A ROLE FOR BAS1 IN THE TRANSCRIPTIONAL RESPONSE TO NA LIMITATION IN *C. glabrata*

by

Carlos Gomez

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

January 2017

Abstract

The budding yeast *Candida glabrata* (an opportunistic pathogen) has emerged as a major health concern, particularly in immunocompromised patients. A key virulence factor Epa1, is the founding member of the Epithelial Adhesin gene family. Structural characterization of the functional N-terminal lectin-binding domain of Epa1 is critical in identifying cognate receptor ligands on host cells. We employed glycan microarray analysis and pursued crystallographic analysis as our first approach. Since *C. glabrata* has lost its *de novo* NAD^+ biosynthetic pathway, limitation for nicotinic acid (NA) or related vitamins in the media leads to NAD^+ depletion in the cell and virulence effects. The known NAD^+ dependent regulators Hst1 or Sir2 do not mediate the virulence effects of NAD^+ limitation.

We previously employed microarrays to analyze the transcriptional response to NAD^+ limitation. Among the several hundred genes induced by NA limitation are large sets of strongly up regulated genes that are orthologues of *S. cerevisiae* genes regulated by the transcription factors Bas1 and Bas2. In *S. cerevisiae*, *BAS1* and *BAS2* regulate several interconnected metabolic pathways, including histidine, purine biosynthesis, tetrahydrofolate metabolism, and phosphate utilization. We

have analyzed the transcriptional response of the Bas1 and Bas2 regulons in conjunction with the Gcn4 regulon in *C. glabrata* in response to NAD^+ limitation. We show that genes in the purine biosynthetic pathway are strongly up regulated by NAD^+ limitation and that this requires Bas1 and Gcn4. We also show that mutants lacking *BAS1* or both *BAS1* and *GCN4* are strongly compromised in virulence consistent with a role during infection, and potentially in mediating the hyper virulence associated with NAD^+ limitation. Furthermore, an analysis of the purine biosynthetic pathway suggest that the effect of NAD^+ limitation on transcription of purine biosynthesis genes may be due to decreased activity of IMP Dehydrogenase (IMPDH) which catalyzes the conversion of inosinate (IMP) to xanthylate (XMP). We present preliminary and compelling results that reveal NAD^+ limitation leads to depletion of guanine nucleotides and to an overall imbalance of the nucleotide pools, which has wide-ranging repercussions. Inhibition of IMPDH in *C. glabrata* mimics some aspects of NAD^+ limitation, including the increase of virulence.

Thesis advisor:	Dr. Brendan Cormack
Second reader:	Dr. Sandra Gabelli
Other readers:	Dr. Susan Michaelis
	Dr. Peter Espenshade

Acknowledgements

I would like to thank my advisor Dr. Brendan Cormack for his ongoing support and advisement; my thesis committee members Drs. Susan Michaelis (chair), Sandra Gabelli, and Peter Espenshade. I would also like to acknowledge past students and postdocs in our lab for all the helpful discussions and support throughout the years. I would like to thank especially Dr. Sandra Gabelli for her dedication and mentorship especially for her guidance provided for data presented in Chapter 2 and beyond. My deepest gratitude goes to my family for their unwavering support and patience; and to my friends for their love, friendship and support. I also wish to acknowledge Drs. Pete Stavropoulos and Sandro Santagata for their friendship and all their helpful guidance.

Table of Contents

Chapter 1 Background

There are approximately 1.5 million different fungal species identified but only about 300 of those are known to cause illness in people. Candidiasis is caused by about 20 different *Candida* spp. which are the most fungal pathogens causing bloodstream infections with mortality rates of up to 60%.[1] While *C. albicans* is the most frequent pathogen, non-albicans Candida species are also important, particularly *C. glabrata*, which is the second most common Candida species causing bloodstream infections. This emerging yeast pathogen is characterized by unique features that warrant considerable attention to manage its diagnosis and treatment. In the general public *C. glabrata* exists as a commensal organism and is detected in about 20-30% of individuals. It is associated with increased morbidity and mortality in immunocompromised patients, including those with AIDS, in cancer patients and in organ transplant patients receiving chemotherapy or immunosuppressive drugs.[2] As a causative agent of candidiasis *C. glabrata* exerts its biggest impact on the elderly population and those passing through hospital ICUs. Age and exposure to antibiotics especially azole type drugs are key indicators for *C. glabrata* infection. The resulting complicated nosocomial infections or UTIs are a consequence of some of the unique features that sets apart this pathogen. It possesses an inherent resistance to azole class anti-fungals. Other virulence attributes

include a family of adhesin genes (Epithelial Adhesin or *EPA* genes) that allow for colonization of epithelial and endothelial cells.

Our work builds on the efforts of previous and current lab members that showed how this EPA family is transcriptionally regulated/repressed. Of particular interest is the role of NAD^+ as a regulator of virulence. The NAD^+ status of *C. glabrata* cells is a key determinant of disease progression. *C. glabrata* is an NAD^+ auxotroph, and cells that are starved for NAD^+ overexpress the EPA family adhesin genes, and are hyper virulent. NAD^+ is critical for signal transduction and a number of metabolic pathways; it is recycled during redox reactions and it is consumed when it functions as an enzymatic co-factor. These subtle features are clinically important since they appear to provide *C. glabrata* a mechanism toward pathogenicity and are clearly demonstrated by its capacity to adhere to abiotic surfaces, notably through indwelling catheters with high incidence rates in hospital ICUs. One of our studies looked to uncover the downstream sensors or pathways that are impacted by an NAD^+ limiting environment which resultantly led to the possibility that GMP or nucleotide imbalances are the true drivers of *C. glabrata* virulence.

Incredibly, since the 1980's, no new classes of antifungals have made it to the clinical stage or beyond with the exception of echinocandin and

its various derivatives. Understandably, the shared eukaryotic nature of fungal and mammalian cells demands the creation of antifungal agents with a high degree of selectivity and specificity for fungal cells. The cell wall is a unique organelle that has proven to be an effective target for developing antibiotics against bacterial infections. A similar strategy has been applied against yeast infections where the major targets are chitin, mannan and (1,3) β-D glucan synthesis. These polysaccharide polymers are absent from mammalian cells; they therefore represent attractive sites of intervention for drug development. Echinocandin lipopeptides are fungicidal with low toxicity and with very low incidence of resistance in *C. glabrata* reported. The drawback with this class of medicines is their poor oral bioavailability hence they are usually only administered intravenously. One current drug trying to address this clinical need is the semi-synthetic enfumafungin derivative MK-3118, originally identified by Merck now under clinical trials by Scynexis known as SCY-078, a novel oral and intravenous drug that has shown to be effective against invasive fungal infections, at least in animal studies.[3]

Our research on the *EPA* family of adhesins, and the initial binding step of infection could possibly lead to the development of a new class of antifungals. In addition, the studies surrounding this family of adhesins culminated in the identification of novel glycan ligands recognized by Epa family adhesins. The results may help identify potential new cell types targeted during infection.

The following work focused on the general mechanisms of *C. glabrata* virulence with the intent of furthering our understanding of specific components of the pathogen's cell wall as well as the identification of regulatory mechanisms governing the response of *C. glabrata* to NAD^+ limitation. Increased advances in our knowledge may ultimately lead to effective new agents against *C. glabrata* or even other invasive Candida spp.

Chapter 2 EPA Lectins and Glycan Ligands

2.1 Introduction

Amongst clinically relevant Candida species, *C. glabrata* is the second most commonly isolated species after *C. albicans*.[4] *Candida glabrata* is an increasingly relevant pathogenic yeast accounting for both mucosal and blood stream infections in humans. The initial step in infection is thought to be adherence of the organism to target human tissues. In the past several years adhesins encoded by the *EPA* gene family have been shown to be the key mediators of the host-pathogen interaction. The founding member *EPA1* (Epithelial Adhesin -1) was discovered to bind avidly to human epithelial cells in culture.[5] Early work also suggested that Epa1 was a Ca^{2+} dependent lectin similar to the *Saccharomyces cerevisiae* Flo1, a surface protein known to be involved in protein – glycan recognition during yeast flocculation.[6] Dissection of the Epa1 functional domains defined characteristic signaling sequences that direct the protein to traffic and localize to the cell wall specifically. Further analysis through sequence homology allowed the identification of additional *EPA* gene family members.[7–9] Over the past several years, our lab (Cormack-JHU) has focused on understanding the adhesion mechanisms and characteristics of the *EPA* family of cell wall localized adhesins discovered in *C. glabrata*, specifically for the BG2-strain.

Most of the known fungal adhesins are GPI - anchored cell wall proteins. They possess a very well characterized modular architecture composed of an N-terminal carbohydrate ligand binding domain followed by a central highly repetitive and heavily O-glycosylated Ser/Thr rich region and a terminal hydrophobic sequence denoting the GPI anchor site.[10] The functional region, located at the N-terminus, in Epa proteins and the Flo family homolog in *S. cerevisiae* has recently been identified as the "PA-14 domain".[11] This lectin domain was originally identified by homology to the 14kDa fragment of the Protective Antigen toxin component of *Bacillus anthracis*. A PSI-BLAST search analysis revealed the presence of a ~150 amino acid sequence as a new β-barrel domain that utilizes Ca^{2+} ions and whose presence has been uncovered in a variety of bacterial and eukaryotic glycosidases, glycosyl transferases, amidases, proteases, bacterial toxins, proteins involved in cell adhesion and in human polycystic kidney and hepatic diseases.[11] The phylogenetic distribution of this motif encompasses a variety of bacterial and eukaryotic species but does not include Archae. This new domain was also found present in the pro-peptide fragment (PA_{20}); a component of the anthrax toxin whose structure has been published.[12] Although related, the critical loops for binding carbohydrates are of different length in the Epa adhesins and presumably impart unique binding selectivity.

In *Saccharomyces cerevisiae*, gene silencing at the telomere is initiated by recruitment of the Sir complex (Sir2, Sir3, and Sir4) to telomeric repeats, followed by spreading into the adjacent subtelomeric region. The spread of silencing is dependent on the function of the sirtuin Sir2, a NAD^+ -dependent histone deacetylase (HDAC). The same mechanism is found in *C. glabrata*. Sir2 becomes inactive in the absence of NAD^+. Since *C. glabrata* is an NAD^+ auxotroph, Sir2 and the related sirtuin Hst1 become inactive in environments limiting for vitamin B3 (Nicotinic acid, nicotinamide, nicotinamide riboside). Under low NAD^+ conditions (the subject of study for chapter II of this thesis) the Hst1 and Sir2 proteins become inactive, consequently gene silencing along the region is relieved. The *EPA* genes are encoded largely in sub-telomeric clusters, subject to Sir2-dependent transcriptional silencing and are subject to a novel form of pathogenic gene regulation whereby transcriptionally-silenced *EPA* genes are de-repressed during infection in host environments limiting for vitamin precursors of NAD^+.[9–16] In our particular lab strain BG2, *C. glabrata* encodes 25 *EPA* genes, 22 of which are encoded in sub-telomeric regions of the genome.[17,18] Three genes in particular *EPA1*, *EPA6* and *EPA7* have been carefully studied and at least for *EPA6* and *EPA7* it has been shown that their expression during a murine model of urinary tract infection (UTI) is likely due to derepression of silencing due to the lack of nicotinamide adenine dinucleotide (NAD^+).[13,16]

Interestingly, the three family members (Epa1, Epa6, Epa7) share >90% identity at the protein level, yet an analysis of their sugar binding specificities revealed a varied carbohydrate binding repertoire.[19] Glycan microarray analysis has been used to determine the sugar binding range of Epa1, Epa6 and Epa7 using both purified proteins and whole *S. cerevisiae* cells. The essential results indicated that all three proteins bind to a terminal galactose residue with notable differences in the binding selections between Epa6 and Epa7 which share the highest sequence (protein) homology in the set.[19] Moreover, Epa1 and Epa7 prefer terminal disaccharides with a β-type linkage while Epa6 shows no preference for either linkage class (α or β).[19] The binding differences between Epa6 and Epa7 have been ascribed to one or two amino acid changes within a small stretch of 5 residues. These results are striking and help explain the binding capacity difference for endothelial cells between these two lectins. Our current studies will support and/or emphasize earlier findings while extending the results to include a new set of *EPA*s namely, Epa12, Epa23 and Epa25. This ongoing study thus far is in keeping with previous overall results for Epa1, Epa6 and Epa7 and presents compelling results for the new set of Epa proteins that show a predilection for more exotic glycan types such as sulfated or sialylated glycans.

Many reasons necessitate a physical understanding at the structural level of Epa proteins that could reveal glycan specificity as well as help clarify the apparent overlapping and redundant functions of the Epa family of proteins. Structural characterization of the functional N-terminal lectin-binding domain is critical in identifying cognate receptor ligands on host cells.

The importance of this family of adhesins as one of the main virulence factors for *C. glabrata* has been firmly established.[18,20] The continued rise of *C. glabrata* infections and the emergence of other non-albicans pathogenic species (NAC) prompts a rigorous study of the basic infectious strategies that could yield important information for prophylaxis protocols as well as new treatment regiments against mycotic infections.

The following studies were undertaken to try and determine the structural basis of Epa-mediated binding to host glycoconjugates and to establish a more comprehensive inventory of the full complement of *EPA* – glycan determinants. By way of structure analysis of the lectin domain of one of the Epa proteins and usage of glycan-microarray (glycochips) screening, we hoped to expand our understanding of the adherence step in the establishment of fungal infections by *C. glabrata*.

2.2 Results

2.2.1 EPA 1 Sequence Analysis and Expression in E. coli

Factors affecting protein expression have been intensely studied to improve on heterologous recombinant protein expression.[21] One of the important determining factors is the influence on translational efficiency of specific amino acids at the +2 position.[22] To this end the Epa1 NT sequence region was put through the secondary structure prediction program - PredictProtein (PP; https://www.predictprotein.org Figure 2.1). Those results along with being guided by the recently described homologous region PA-14 domain and the known crystal structure of anthrax toxin allowed us to narrow down the size of our starting functional Epa N-terminal domain. This initial analysis aided in the production of a panel of expression constructs that ranged in size between 17 kDa and 35 kDa and in lengths of 150-320 amino acids as detailed in Table 2. A general N-end rule for predicting enhanced protein expression is the incorporation of Ala, Ser or Leu after the start Met residue whenever possible.[23] The prediction program depicts a heavily unordered loop structured domain (L=67%) along with several β-sheets (E=33%) shown in Figure 2.1. The above considerations are certainly important but were rendered moot, once we decided to express these fragments as N-terminal fusions to His-Sumo1.

The cloning schematic for Epa1-NT protein fragments is depicted in Figure 2.2 and further described in the Materials and Methods section. Sumo-1 has been well characterized and established in aiding in the solubility of a large number of recombinantly expressed proteins ultimately used for structural studies. Coomassie stained SDS-PAGE gels confirm the heterologous expression of our set of 14 fragments as His-Sumo fusion proteins. The results from small Ni^{2+}-NTA resin batch purification experiments shown in Figure 2.3 indicate an abundance of presumed degradation products. Fractions corresponding to elutions and resin are shown in panels A and B. Although there is substantial expression of practically all protein fragments there is marked protein instability as indicated by the large quantities of degradation products and/or aggregated protein near the lower molecular weight range of the gels (20-25 kDa). Nonetheless, PAGE migration analyses as evidenced on panels C, D and E from samples taken before and after IPTG induction do show the expected variations in protein sizes reflecting all the different constructs that were tested. Even so, the elutions samples show apparently unstable and improperly folded protein products.

There are several possibilities that could account for the disappointing behavior of all of the constructs. The His-sumo fusion tag could itself be interfering with the dynamics of the initial folding process. Many instances exist in the literature demonstrate that proteins do

benefit from a change in the placement of the affinity tag so a C-terminal tag rather than an N-terminal position could allow for better folding outcomes. In an attempt to improve protein solubility and stability we tested a number of parameters: salt, pH, reducing capacity and detergent solubility (Figure 2.4). The different buffer formulations were introduced, as crude cell lysates were prepared. Epa1 (version 1.2) overexpressed in bacterial cell pellets were subjected to various experimental buffers and carried through to the purification step in Imidazole elutions, yet none of the 11 different conditions tested from Table 2.2 did much to prevent protein instability.

Apart from the physical parameters that were addressed we considered more importantly the inherent attributes associated with the *EPA* family of lectins. Of real consequence are the facts that this N-terminal domain is i) rich in Cysteine residues, with 5 predicted S=S bonds and ii) it possesses predicted post-translational N-glycosylation modification sites along residues (**N^{140}** and **N^{333}**). Several potential N-myristoylation sites were also predicted but likely do not affect expression. Given these features, it seems likely that the problem with protein stability has more to do with implementing the appropriate expression system rather than the conditions during the purification process. Consequently, we decided that a eukaryotic approach might be more accommodating for the particular characteristics of these lectin

proteins. Although currently there are new bacterial systems that are suited to expression of glycosylated and disulfide bond rich-proteins we opted to express the Epa domains in a eukaryotic heterologous expression system.

2.2.2 Expression of Epa 1 in mammalian cells

The DNA fragments encoding Epa1 were transferred from our bacterial expression plasmids and subcloned into a mammalian expression system for production of secreted proteins in adherent HEK 293 ($GNTI^-$) cells. The cloning and purification protocols are detailed in the Materials and Methods section. Twenty construct fragments were created for metal chelation purification.[21] Small batch experiments using 6-well plates yielded mostly soluble, non-degraded proteins as shown in Coomassie stained PAGE gels (Figure 2.5). We cannot make any definitive conclusions with regards to the influence the various C-terminal positions had on the stability of the panel of constructs. An overview of panel A - the elution fractions, suggests there could be an ideal fragment length imparting a stability influence that is exerted by different C-terminal positions, at least for the **A^{30}** group of constructs. Interestingly, we see that there is a complete loss of protein expression as the fragment becomes progressively more truncated from the N-terminus. Inspection of the primary sequence and our original secondary structure analysis shows that there exists a **Cyc^{50}** that is predicted to be part of a

disulfide bridge. Notably, the two sets of fragments where expression was successful started at A^{30} and L^{38}. The next group, initiating at S^{62} failed to express and shows that **Cys^{50}** is critical for the overall architecture of this lectin domain. In fact, the solved structure has borne out the fact that Cys^{50} along with Cys^{179} help tether the N-terminal end to a loop between β6 and β7 rendering the desired compact shape for the protein.[24] This Cys^{50} is also preserved in the distantly related Flo5 protein from *S. cerevisiae,* where the PA14 domain has also been identified - suggesting a critical role for this residue. The lower panel B shows that there still remains a good percentage (~30-40%) of aggregated protein bound to the Ni^{2+} resin as well as some unknown non-specific Ni^{2+} binding protein.

In an effort to arrive at a more stable and well-expressed protein, a limited proteolysis experiment was conducted. Construct Epa 1.5 was subjected to several different concentrations of Proteinase K and visualized on a PAGE gel after Coomassie staining (Figure 2.6 insert). After overdigestion the sample was further processed for mass spectrometry analysis (Figure 2.6). These analyses helped identify an extremely stable core domain bounded by **Ser^{40}** and **Ile^{283}**. Once this core domain was uncovered, new sets of constructs were sub-cloned using the corresponding residues for Epa 6 and Epa 7, as listed in Table 2.4. Small-scale expression tests were confirmed by Western analysis

using an N-terminal specific antibody (lab-stock #3638) as displayed in Figure 2.7. Epa-NT mammalian expression is known to be low yielding showing only faint banding in Coomassie stained gels that are poorly visible in the Figure 2.7 image. Of note is the observation that although Epa1, Epa6 and Epa7 are of the same 1° sequence length there is a clear marked difference in the migration pattern displayed on the Western. GNTI$^-$ cells are deficient only for complex glycosylation structures but remain competent for simple mannose modifications. A reasonable assessment is that there are glycosylation differences between these genes although they share over 90% homology at the amino acid level. The functional consequences of these features remain to be fully delineated and explored.

2.2.3 Purification of Mammalian Expressed Epa1p

A two-step purification scheme was developed for the Epa1. As detailed in the methods section, Epa1 was harvested as a secreted protein from HEK-293 (GNTI$^-$) cells and passed over a Ni^{2+} IMAC column, followed by Ulp1 enzymatic treatment overnight for His-Sumo1 fusion removal. The collected protein (FT; flow-thru) was subsequently purified over a Lactose-sepharose column yielding a highly purified Epa1 species fraction (Figure 2.8 – Coomassie stained insert). Samples collected throughout each of the two purification steps were collected for analysis on SDS-PAGE. The Coomassie stained gel shows practically no or

minimal protein being captured by the resin in an aggregated manner. A portion of this material was then analyzed by HPLC using an S75 diagnostic gel filtration column (Figure 2.8). The single peak UV trace offers a good indication that the protein is of a single mono-disperse species. Additional information about the general behavior of the protein can be inferred by the lack of detectable protein in the void volume; the latter is a sign of the presence of misfolded or aggregated proteins.

Marginally better protein yields from Epa7p prompted an initial full-scale production for protein crystallization experiments. As detailed in the mammalian expression methods – once Epa7 was purified and concentrated (6 mg/ml), the protein was screened for crystallization conditions utilizing the Mosquito LCP robot on 96-well plates. Several commercial screens were attempted such as Index- (Hampton Research) and JCSG+ (Qiagen), ultimately the latter of which resulted in visible micro-crystals of Epa7 (Figure 2.11, A). Epa7 crystals grown by hanging drop vapor diffusion are shown in panel A (Figure 2.11) and verified as from protein of interest through Western-blot analysis panel B. To remove the ambiguity of protein vs. salt crystals a couple of crystals were gathered for confirmative testing using our Epa1 N-terminal specific antibody (#3638) in panel B (Figure 2.11). Regrettably, the crystals were tested for X-ray diffraction but yielded poor resolution data (about ~ 7Aº). Encouraged by our progress, optimization experiments were underway to

obtain properly packed protein crystals. However, during this time the protein structure of Epa 1 was published[24] by Ielasi et.al. 2012 and the structural biology part of the project was stopped.

2.2.4 Purification of EPA gene family proteins

The Cormack lab has previously shown the adhesin ligand specificities of the adhesin proteins Epa1, Epa6 and Epa7 through cell-based experiments using fusions to the C-term domain of Cwp2[19] a GPI-CWP that is important for the membrane structural stability of *S. cerevisiae*. The N-terminal domains were expressed along the cell surface of intact *S. cerevisiae*.[19] While this heterologous expression system worked well there still remained a lack of ligand binding information from natively expressed proteins. We sought to determine the glycan binding specificities through the expression of the entire *EPA* gene family. Of the 25 gene members 20 constructs of the N-terminal domain were successfully created and subcloned using restriction enzymes *Bam*HI and *Xho*I for directional cloning or compatible *Bgl*II and *Sal*I fragments (detailed in Table 2.4). The exact cloning regions were determined after a multi-sequence alignment for all *EPA* gene members was created. The alignment was made using the open resource T-coffee (http://tcoffee.crg.cat). Conspicuously, the alignment reveals the potential presence of the lectin binding domain known to be present in Epa1, Epa6 and Epa7 may also be present in several Epa protein family

members (Figure 2.10). Eleven of the 20 proteins were expressed as 6xHis-Sumo fusions and secreted from HEK 293 (GNTI$^-$) cells. A summary of the expressed proteins is shown in Coomassie stained gels (Figure 2.9). All protein fragments were made as His-Sumo1 fusions with the exception of Epa1, Epa6 and Epa7 (not shown), which are additional versions possessing a simple 6xHis tag. We ultimately obtained glycan specificities for 6 gene family members, namely *EPA*s: 1, 6, 7, 12, 23, and 25. These results are summarized on Table 2.5 – Glycan Microarrays.

2.2.5 Glycan microarray analysis (EPA1, EPA6 and EPA7)

With 50% or more of all proteins carrying glycan chains, glycomics has emerged as an area of intense investigation in the post-genomics era.[25] Glycan microarray analysis is a high-throughput approach for determining the ligand-binding specificities of a range of glycan binding proteins (lectins). For this current work we used printed glycan arrays developed by the Consortium for Functional Glycomics (CFG; http://www.functionalglycomics.org), to screen purified proteins of the *EPA* gene family. Glycan specificities were obtained for the following *EPA*s: 1, 6, 7, 12, 23 and 25 using two different protein concentrations (50 µg/mL and 200 µg/mL). Prior to analysis the proteins were fluorescently labeled by conjugation with the Alexa-Fluor 488 dye for detection. The printed arrays contained both natural and synthetic

carbohydrate structures for a total of 612 unique glycans representing the major glycoprotein and glycolipid structures.

Our analysis show that all three recombinantly expressed lectin domain proteins: Epa1, Epa6 and Epa7 all bind to galactose terminal end ligands which is generally in keeping with previous overall results from glycan microarray experiments. For determining the differences between ligand specificities, the non-reducing galactose end sugars were placed into subcategories defining certain features within terminal disaccharides such as: type of anomeric conformation, the linkage position of the hydroxyls involved, identity of the penultimate residue, modifications of these residues (i.e. Sialylation, N-acetylation, Sulfation) and/or presence of branched structures.

Our results indicate preferential binding by Epa1 and Epa7 primarily to β-type glycan linkages, which corroborates with previous results from our lab, specifically β1-3 linked galactosides (Figure 2.12 Ai, ii). A new distinction from earlier experiments that showed Epa6 binding equally well to β and α-type linkages, our current analysis shows Epa6 strongly prefers binding to the α-linkage class. Within the β–linked class, another divergence from previous results where no preference between β1-3 or β1-4 linkages was found, our current findings for all three adhesins (Epa1, Epa6 and Epa7) show a strong affinity for β1-3 type linkages. All

three adhesins proteins show a similarly strong predilection for β1-3 linkage, essentially exclusive for β1-3 type bonds (Figure 2.12Aii). According to previous studies within this Galβ1-3 sub-category, there was a strong affinity by Epa1 for GalNac in the penultimate position, 'the T-antigen'. Our current data continues to support this as true, yet the selectivity over the unmodified galactose (GalNAc) is much more pronounced for Epa7 rather than Epa1 (Figure 2.12 Aiii). Interestingly, within this β1-3 type linkage group all three adhesins Epa1, Epa6 and Epa7 show an aversion for GlcNAc in the penultimate position compared to GalNAc (Figure 2.12Aiv). The difference between galactose and glucose is due to the C4-OH orientation; this subtle difference reflects the refined specificity that has been attained by *C. glabrata* adhesins. Epa6 highlights this idea by not only preferring α-type bonds but also possessing the capacity to bind Gal β1-3 type linkages over β1-4 types. Clearly the β1-4 bond presents a binding barrier for all three adhesins and the penultimate residue appears to serve an important role in binding criteria as well. Both Epa1 and Epa7 prefer the penultimate galactose to be N-acetyl modified and when the modification takes place on the penultimate position, the type of residue dictates binding success. All three adhesins bind preferentially to galactose rather than glucose (Galβ1-3GalNAc vs. Galβ1-3GlcNAc).

The linkage orientation between the terminal galactose and the penultimate residue is a key aspect in ligand recognition. As mentioned above, the Gal β1-3 type linkage is preferred over β1-4 types, yet within this β1-4 subcategory we find very strong selectivity characteristics as to the identity of the penultimate residue and whether or not the penultimate residue is modified. Further analysis of the β1-4 subcategory shows that glucose (Galβ1-4Glc) in the penultimate position is robustly preferred over galactose (Galβ1-4Gal) and if the glucose residue is N-acetyl modified than even more so than the unmodified glucose (Figure 2.12Bi,ii, iv). This modification selectivity along the penultimate position (Figure 2.12Biii) is favored even when the identity of the residue is a galactose (Galβ1-4GalNAc). Interestingly, this N-acetyl modification at the penultimate position is an important general binding feature since all three adhesins (Epa1, Epa6, Epa7) bind strongly and selectively compared to the simple unmodified sugar whether it is glucose or galactose (Figure 2.12 Biii, iv). Thus, creating a β1-4 type bond with glucose averts the C4-OH orientation that is inhibiting for a β1-3 glucose type linkage (Galβ1-3Glc). This is unlike what we see with the β1-3 bond where galactose is preferred, which is reasonable considering the axial C4 hydroxyl-orientation of galactose. When assessing the influence of the glycan bonds on N-acetyl modified sugars at the penultimate position we see some clear distinctions. As mentioned previously in a β1-3 conformation the three adhesins will bind to GalNAc compared to

GlcNAc, unless the bond is a β1-4 linkage type, in which case Epa6 and Epa7 will be selective for glucose (GlcNAc); Epa1 maintains a modest affinity for GalNAc (Figure 2.12Bii). This is a sharply different result from previous analysis that report an inability for this set of adhesins to distinguish between Galβ1-4Glc and Galβ1-4GlcNAc. This general binding feature is true also in the β1-3 category for Epa1 and Epa7 although not as pronounced.

All three adhesins appear to require galactose in the penultimate position. Epa1, Epa6 and Epa7 bind exclusively to Gal – Gal or Gal – GalNAc glycans but strictly when the bond is of the β1-3 nature and not the β1-4 linkage type (Figure 2.12Ci,ii). Notably, when the penultimate residue N-acetyl modified and the identity is glucose (Gal – GlcNAc) then the β1-4 linkage is strongly preferred by Epa1 (Figure 2.12 Ciii); Epa6 and Epa7 show no bond type preference when the penultimate residue is glucose (GlcNAc). Therefore, as mentioned earlier, a single OH-reorientation can create a large difference in adhesin ligand recognition.

One particular modification feature on glycans that creates a strong preference or strong inhibition for adhesin binding recognition is the presence of sulfation (S). Selectivity is also influenced by the position and the number of the sulfate marks on different glycan residues. The presence of sulfation appears to dominate within the β1-4 linkage

category, but only Epa1 seems to possess an affinity for this modification although under very specific conditions (Figure 2.12Di-iv). Epa1 will bind to a terminal sulfated-Gal (Figure 2.12Di,ii) but not in the presence of two sulfates ((6S)(3S)Galβ1-4GlcNAc). It will also be inclined to bind a penultimate sulfate mark (Figure 2.12Diii,iv) if the linkage is a β1-3 type; Epa6 will also work if it is a β1-4 bond (Galβ1-4(6S)Glc). The Epa1 binding range is exhibited not only by its capacity to bind (S) glycans on two different sugar types but also on two conformations; β1-3 and β1-4 bonded glycans. Epa6 and Epa7 share a similar binding profile to this glycan but show a clear preference for the non-sulfated GlcNAc. These nuanced binding features can be uncovered continuously, but what we can conclude is that sulfation within this set of adhesins does not play a significant role in binding as it does with Epa12, Epa23 and Epa25, discussed below.

Finally, Epa1 and Epa7 possess virtually no binding capacity for α-type linked glycans, which contrasts completely with Epa6, which shows a strong preference for α-types. All three adhesins show similar binding profiles for various α-types regardless of linkage position or penultimate sugar identity (Galα1-3, Galα1-4), except for Galα1-3fucosyl substituted glycan that is strongly disfavored by both Epa6 and Epa7 compared to Epa1 (Figure 2.12 Civ). This is merely a technical point considering the

small relative fraction of α-type linked glycans to which Epa1 and Epa7 bind but again emphasizes the importance of penultimate modifications.

*2.2.6 Glycan microarray analysis (**EPA**12, **EPA**23 and **EPA**25)*

Our analysis show that all three recombinantly expressed lectin domain proteins: Epa12, Epa23 and Epa25 all bind to galactose terminal end ligands which is generally in keeping with previous results from glycan microarray experiments. For determining the differences between ligand specificities, similar categories and subcategories were compared as the above analysis with Epa1, Epa6 and Epa7.

Unlike Epa1, Epa6 and Epa7, which preferentially bind to β-type linkages, Epa12, Epa23 do not. While Epa25 does prefer β-type links it does also have the capacity to bind α-type glycans. There are some characteristic distinctions that distinguish these adhesins from the prior set of proteins. In general there are more modest binding specific profiles regarding the type of linkage, the bond position (C4 vs. C3), the identity of the penultimate residue and a diminished influence of N-acetyl modified penultimate residues when compared to the above discussed set of adhesins.

Our newest set of Epa proteins certainly have a similar general binding specificity for glycans with terminal galactoses however they

display a predilection toward α-type linkages rather than the β-type conformation, particularly for Epa12 and Epa23 (Figure 2.13Ai). Within the α-type linkage class, all three adhesins (Epa12,23,25) prefer the C-4 hydroxyl position (Figure 2.13Aii), contrary to the C-3 of the above set of adhesins, a likely consequence of steric constraints. When comparing Epa12, Epa23 and Epa25 to the earlier set of adhesins we can see that there must be some general structural features between both sets of proteins that allow for some general ligand binding requirements. This new set of adhesins appears a bit more forgiving as to the type of linkage or the identity of the penultimate sugar. While Epa1, Epa6 and Epa7 maintain an exclusive affinity for β type links, specifically β1-3 linkages, Epa12 and Epa23 likewise prefer β1-3 linkages yet possess the capacity for binding β1-4 types as well (Figure 2.13Aiii). Additionally, within this subclass (β1-4) all three adhesins show a preference for a N-acetylated penultimate glucose over the unmodified sugar; Galβ1-4GlcNAc vs. Galβ1-4Glc (Figure 2.13Aiv).

Within the β-type category the nature of the glycosidic bond and the identity of the penultimate residue appears to be most critical for Epa25 binding. Modifications of the penultimate residue as opposed to the terminal galactose also seem to be a distinguishing selection characteristic for Epa25 binding as it possess the most selective binding range than does Epa12 and Epa23. Furthermore, Epa25 shares a similar

binding profile as the above Epa7 adhesin, in that they both have a unique preference for the 'T – antigen' Galβ1-3GalNac (data not shown). These ligand binding distinctions regarding linkage types and C-positions reveal significant structural differences between this current set of adhesins and the former set (Epa1, Epa6, Epa7) that are clearly borne out by the strong preference for sulfated modifications along the ultimate and penultimate glycan positions. Interestingly, while Epa6 also prefers α-type linkages it can only bind preferentially to sulfated modifications under specific circumstances; when the sulfate is on the penultimate glucose residue.

In sharp contrast to the above set of adhesins where only Epa1 shows strong binding to terminally sulfated glycans Epa12, Epa23 and Epa25 all show strong, avid binding for sulfated terminal galactoses (6S)Galβ1-4Glc+/-NAc (Figure 2.13 Bi-iv). When a second modification adds an additional negative charge to the terminal galactose by way of sulfation this becomes inhibitory for Epa25 binding (6S)(3S)Galβ1-4GlcNAc, however, it is clearly preferred by Epa12 and Epa23. In addition to showing a strong propensity to bind to terminal galactose residues, all three *EPA* genes (Epa12, Epa23, Epa25) also bind preferentially to the penultimate (S) modified glucose (Figure 2.13 Biii). Moreover, sulfation at both the terminal (6S)Gal or penultimate residue GlcNAc(6S) is preferred by all three adhesins Epa12, Epa23, Epa25.

Therefore, the capacity to bind glycans sulfated at either the terminal or penultimate residue indicates that it is likely that these adhesins share some common structural binding motif that allows for these negative sulfate groups to be accommodated at either position. One exception to the overall binding preference to sulfated glycans by this set of adhesins is Epa12, which shows strong aversion for the penultimate (S) modified GlcNAc though only when it occurs in the β1-3 configuration (data not shown). Finally, an analysis of the α-type binding category shows that Epa12, Epa23 and Epa25 prefer binding to α1-4 type linkages, unlike the previous set of adhesins that prefer β-type bonds.

The most compelling result from the following glycan-array analysis is the fact that this new set of adhesins analyzed (Epa12, Epa23 and Epa25), clearly have a strong bias for sulfated glycans. And the (S) modification can occur at the terminal galactose, the penultimate residue or both with either a single (S) modification or even double (S) modifications. This is a keen distinction from earlier adhesins analyzed (Epa1, Epa6 and Epa7) where ligand binding to sulfated glycans occurs only under a very limited number of circumstances.

The results highlight and emphasize previous findings using heterologous expression of Epa1, Epa6 and Epa7 in whole *S. cerevisiae* cells. Some discordant findings are somewhat expected considering the

increased range of glycan structures probed in the current arrays pertaining to these studies, that is, the present arrays contain twice the number of glycans used in previous glycan-chip profiling experiments Table 2.5.

2.3 Discussion

C. glabrata survives as a commensal organism in humans until the appropriate conditions present themselves whereby its 'Jekyll and Hyde' persona is exposed and assumes an opportunistic pathogenic existence. The mechanism of adhesion is based on several important factors. The type of proteins expressed by the fungal pathogen, the type of host cells present on the organism and the conditions presently that will be permissible for colonization. The first step in fungal colonization and invasion of host tissues is the adherence of Candida spp. through a set of proteins collectively known as adhesins. A proteomics and bioinformatics analysis of the *C. glabrata* genome has revealed a set of 67 adhesin-like proteins of which 2/3 are subject to subtelomeric regulation. Depending on the *C. glabrata* strain and nutrient sufficiencies there will be 23 to 26 gene members held in a transcriptionally repressive state. These are the primary mediators of adhesion belonging to the *EPA* (epithelial adhesin) family of proteins, the largest of seven subgroups identified. Much of the success of *C. glabrata* as a pathogen is owed to its repertoire of adhesins and to the inherent capacity of expansion and allele variability within the

adhesin gene families that is likely created during DNA recombination events. Adhesins are known as virulence factors that contribute significantly to the pathogenicity of *C. glabrata* as well as *C. albicans* (Als protein family).

Identification of a highly stable protein fragment within the lectin-binding region of Epa1, Epa6 and Epa7 (S40-I283, W40-V283 respectively) was revealed during the cloning and expression optimization experiments. These early results imparted an initial indication of the importance of disulfide bridges to the overall stability and architectural integrity of this N-terminal domain. Recent data from the solved crystal structure of Epa1 indicate that both the N-terminus and C-terminus are bound to the main body of the protein by disulfide bridges, one between Cys^{50} and Cys^{179} and the second between Cys^{180} and Cys^{262}. This arrangement gives the protein a more compact shape and blocks access to the space between two central sheets thereby decreasing the surface of the β-sandwich exposed to the solvent. The overall structure of the barrel-shaped Epa1-NT domain is of an antiparallel β-sandwich motif composed of 15 β-strands, 11 of which encompass the PA14 domain. The remaining 4 β-strands along with the termini wrap around the sandwich and help limit access to solvent. The reported binding site is mediated by a series of 5 loops that utilize a unique D*cis*D motif to help coordinate Ca^{2+} for direct interaction with hydroxyl groups essential for glycan

specific binding. Another critical disulfide bridge crosslinking two loop fragments (Cys^{78} and Cys^{119}) appears to be very important in shielding the Epa1-NT active site from the external environment necessary for efficient and specific binding to carbohydrate ligands. A similar layout is also employed by the *S. cerevisiae* homolog Flo5A, whereby a unique subdomain together with a second loop help seal the surface of the underlying β-sheet from solvent access[24,26,27]. In addition to limiting solvent access to the active site another characteristic afforded by these cysteine linkages may be to provide rigidity to the overall protein to facilitate binding along the extracellular surface of the cell. This conserved shared strategy suggests a general feature that may be significant for GPI linked cell-wall-protein adhesins.

Results from glycan-array profiling experiments provide further evidence behind the nature of Epa-specific binding to host epithelial and endothelial cell types. The overall results are consistent with previous reports and help emphasize the importance of determining the glycan specificities for all members of the *EPA* gene family. One of the most pronounced outcomes among all adhesins studied to date is the overwhelming preferential binding of Epa1 and Epa7 for β-linked galactosides over α-type linkages. The αGal epitope (Galα1-3Gal) modified glycoproteins and glycolipids can be found in all mammals except humans, apes and old world monkeys[28,29]. One straightforward

explanation offered is that, evolutionary processes have resulted in loss of the gene necessary for expressing these αGal markers on their cells. In turn, *C. glabrata* has evolved along with humans to develop sugar-binding specificities that delineate its environment. It would be of some interest to compare *C. glabrata* strains in animals with the strains found in humans to see how the *EPA* family of adhesins has evolved if at all to any significant extent. Furthermore, of physiological relevance is the preference of Epa1, Epa7 and Epa25 for the T-antigen (Thomsen-Friedenreich), Galβ1-3GalNac glycan. It is a main core structure found in glycoconjugates of mucin type proteins – a high molecular weight family of proteins that are secreted by epithelial cells as principal components of mucus. Yet another sugar moiety found abundantly along the mammalian digestive tract is the glycan GlcNacβ1-3GalNac that is curiously of low affinity to the Epa adhesins analyzed herein. With the exception of Epa1, which showed marginal capacity for terminal GlcNacs (data not shown), all of the Epa proteins tested in this study displayed a strong predilection for terminal galactose glycosides. Binding analysis of these adhesins allows for a much sharper picture of the manner in which *C. glabrata* is able to survive as part of the human gut and/or exploit characteristics associated with constituents of the microbiota.

Contrary to the previously analyzed set of *EPA* genes, the recent glycan specificity data for Epa12, Epa23 and Epa25 precedes an

extensive body of biological data. Apart from potential adhesion redundancy, the physiological roles for these *EPA*s are unknown. In like manner, a fundamental understanding of the regulatory mechanisms controlling the expression of each of these family members remains to be fully established. Cell binding data produced from our lab has established some initial indications as possible cellular adhesins; yet much work remains to be developed to arrive at a contextual framework for the functional mechanisms for this set of *EPA* genes.

Our results point to the fact that it is of some consequence that the Epa12, Epa23 and Epa25 are especially inclined to bind to sulfated terminal disaccharides. Sulfated glycans play important roles in biological function such as homing of lymphocytes[30,31] and adhesion of neural cells.[32] Heparin – a heavily sulfated GAG (glycosaminoglycans) is known to be one of the most negatively charged polymers of natural occurrence, an essential feature allowing for its participation in many molecular and cellular events in inflammation.[33] Recognizing the general localization of the *EPA* gene family – one conceivable interpretation is that evolution of *C. glabrata* throughout its human residency has rendered a subset of genes (*EPA*s: 12, 23, 25) studied thus far with the capacity to co-opt a similar strategy utilized by human leukocytes for obtaining access to endothelial cell types at sites of inflammation. The routine use of clinical catheters is intimately associated with chronic

activation of the immune and inflammatory system.[34] This work suggests Epa-mediated binding to sulfated glycans may provide a link between *C. glabrata* UTIs, catheterization and candiduria, in hospitalized patients,[35,36] in whom indwelling catheters are a very important risk factor.

Sulfated polysaccharides are actively being analyzed as drug candidates for protozoan infectious diseases,[37,38] along with sulfated proteoglycans and glycosaminoglycans (GAGs) both of which are known targets of a number of infectious diseases with significant health consequences in immune-competent humans. GAGs are found on the surface of almost all cell types and are used by many intracellular pathogens to attach and gain entry to host cells.[39] Microsporidia are opportunistic protists that infect a wide variety of animals, including humans. *Encephalitozoon intestinalis* exploits sulfated glycans such as the cell surface (GAGs) in selection of and attachment to host cells.[40] Other eukaryotic microbes utilizing a similar targeting strategy like *Plasmodium falciparum* bind heparin and heparin sulfate,[41,42] allowing for the infection of hepatocytes and the placenta. The success of *Toxoplasma gondii* infection is attributed to its ability to bind a wide variety of host GAGs.[43] The bacterial pathogens *Listeria monocytogenes* [44] and *Mycobacterium* spp.[45] also bind cell surface heparin for adherence to epithelial cells. The dengue [46] and foot and mouth disease viruses also

interact with heparan sulfate.[47] The parasites *Trypanosoma cruzi*,[48] *Plasmodium spp.*,[41] and *Leishmania spp.*[49,50] are reported to use heparan sulfate in host cell adhesion. In all, binding to host cell surface sulfated targets is an established adhesion mechanism that might also be employed by specific *EPA* family members of *C. glabrata*.

It should be noted that the binding profiles obtained here through heterologous expression of the core lectin binding domain alone lacks the native structural components that have been shown critical for the function as adhesion proteins. Explicitly, the lack of its heavily O-glycosylated central domain which is known to impart physical rigidity to the N-terminal binding region could be influencing the extent of binding variability in our analysis. A set of complementary cell based experiments utilizing exogenously expressing fusion proteins to the Cwp2 C-terminal domain of *S.cerevisiae* as demonstrated in earlier work should be undertaken to help support current findings.

One final consideration regarding the recently published Epa1 structure concerns the manner in which the protein was generated. The reasons behind our painstaking efforts for developing a eukaryotic expression system was due to the knowledge of the fact that Epa1 undergoes several post-translational modifications, principally – glycosylation. Recent binding affinity differences between the Epa1 lectin

homologs Flo1, Lg-Flo1 and Flo5 exposes the affects of deglycosylation on substrate binding affinities.[51] The latter recombinantly expressed in *Escherichia coli* and the first two in yeast. Thus it remains to be seen if the current structure of Epa1 produced in *E. coli* is physiologically consistent and compatible with binding selectivity of our *EPA* adhesins generated in mammalian cells.

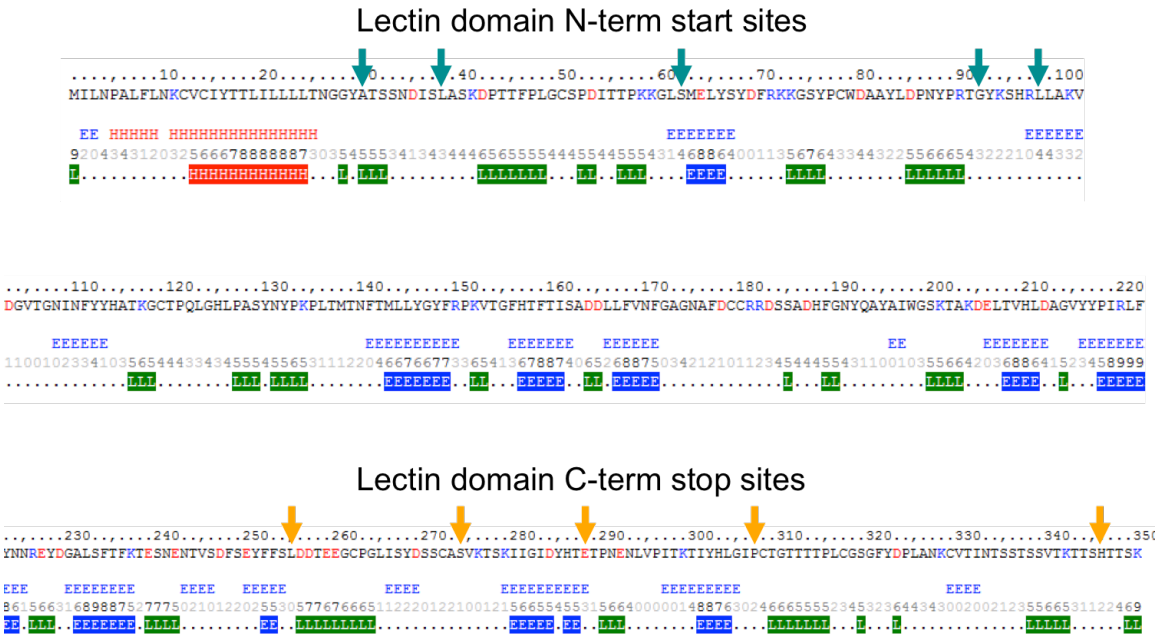


Figure 2.1 Secondary Structure Prediction of C. glabrata EPA1 NT (ATG-K350).

Green arrows indicate N-terminal start sites. Orange arrows indicate C-terminal stops. The single letter amino acid sequence shown is numbered (above) along with probability estimates (below) for secondary structure features. (The more likely the feature the higher the number value 0 – 10). Constructs were designed as indicated by green and orange arrows with a general cloning strategy to avoid the disruption of structured domains such as extended β-sheets and/or unordered loop regions. Specific construct domains are detailed in Table 2.1.

H – α helix, **L - loop**, **E –** β strand.

(https://www.predictprotein.org).

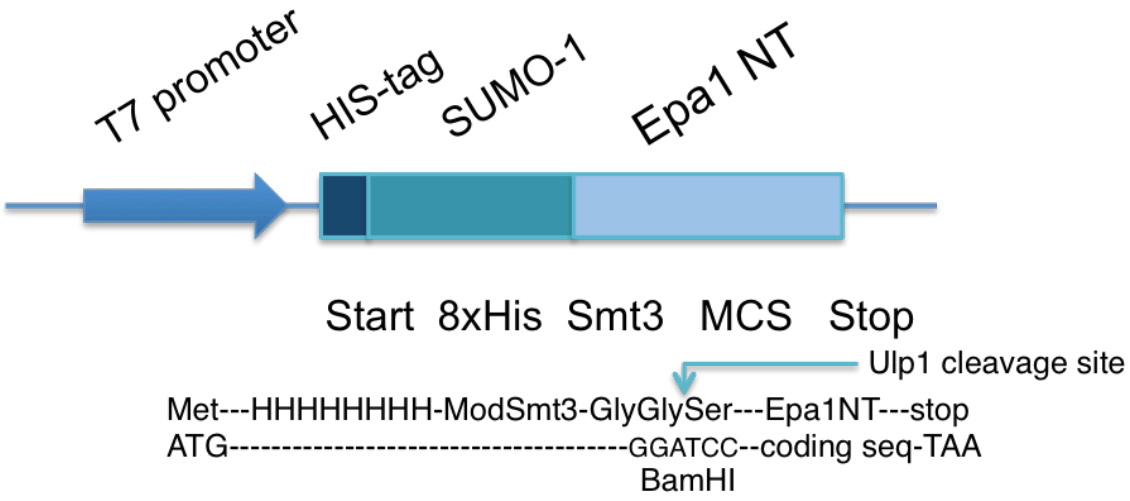


Figure 2.2 Schematic representation of the cloning of His-Sumo-Epa1 NT fusion

Protein expressed in *E. coli*. (BL21-DE3). Epa 1 fragments are immediately downstream of Sumo (Smt3) and designed to be cloned in-frame by specific usage of the *BamHI* restriction site. The system allows for removal of the Sumo protein along with the polyhistidine affinity tag by way of the Ulp 1 (Ubiquitn-like specific protease) protease that specifically recognizes the terminal Gly-Gly sequence of Sumo. After cleavage, the result is the incorporation of a non-native Serine residue 5' to the native sequence.

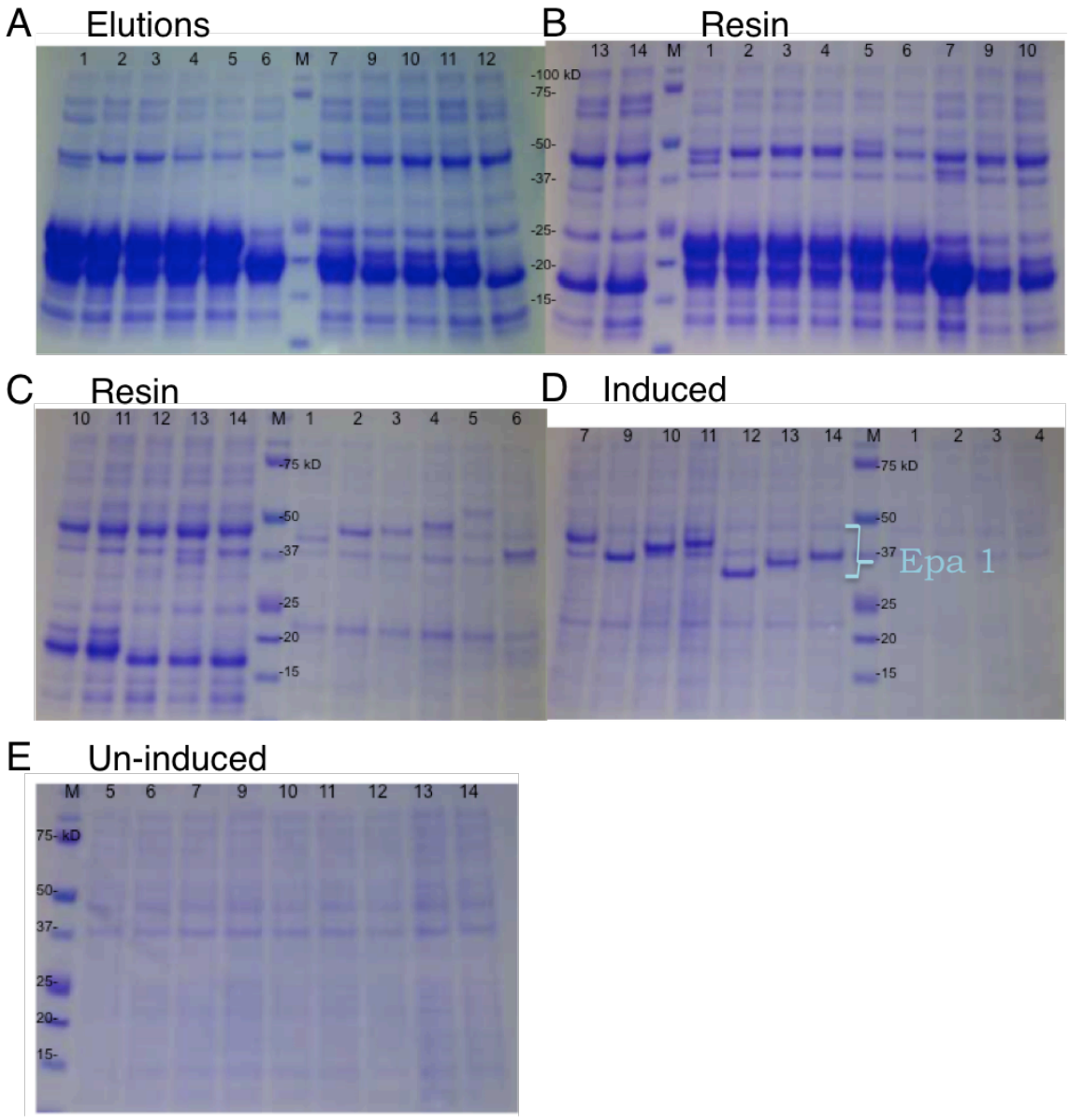


Figure 2.3 SDS PAGE gels of Epa 1 NT fragments.

Coomassie stained gels of a panel of His-sumo Epa1 (N-terminal) fusion proteins. Samples were run from lysates of *E. coli* strains of various stages along the expression tests. Lane numbers correspond to different protein constructs of Epa1 as detailed in Table 2.1.

Blue bracket indicates the range of expected Epa1p fragments sizes (~ 32 – 52 kDa) including the Sumo tag and potential glycosylation.

(M lane = protein standards marker).

A, B. Imidazole Elutions, M, Elutions, M, Resin bound fractions (boiled off in1xLB).

C, D. Resin bound, M, Induced proteins, M, Un-induced fractions.

E. M, Un-induced fractions.

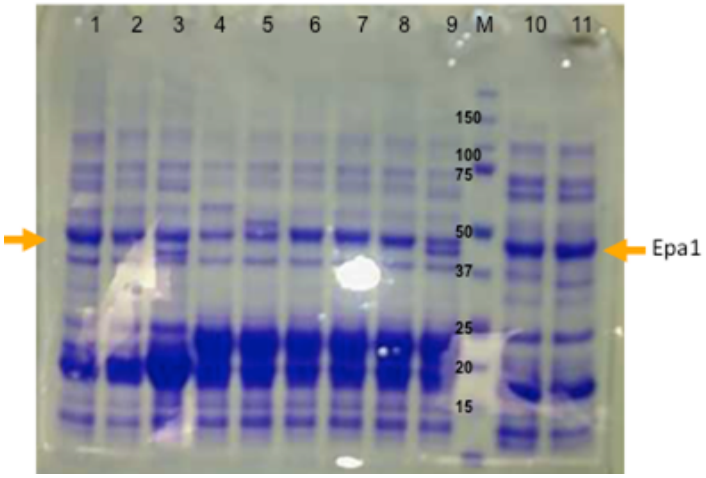


Figure 2.4 Coomassie stained gel of EPA1 construct #2 (A30-S274).

Using the various buffer conditions listed on Table 2.2, samples were Ni^{2+} affinity purified. Numbered lanes correspond to separate buffers used during the purification process. (M – prestained protein molecular weights standards). Small-scale batch purifications were made whereby one large culture (150 mL) was pelleted in 10 mL fractions at which point the testing buffers (Table 2.2) were added and carried thru the affinity purification process. No clear benefit was afforded. Protein appears to express (indicated by orange arrows) but is mostly degraded low MW products.

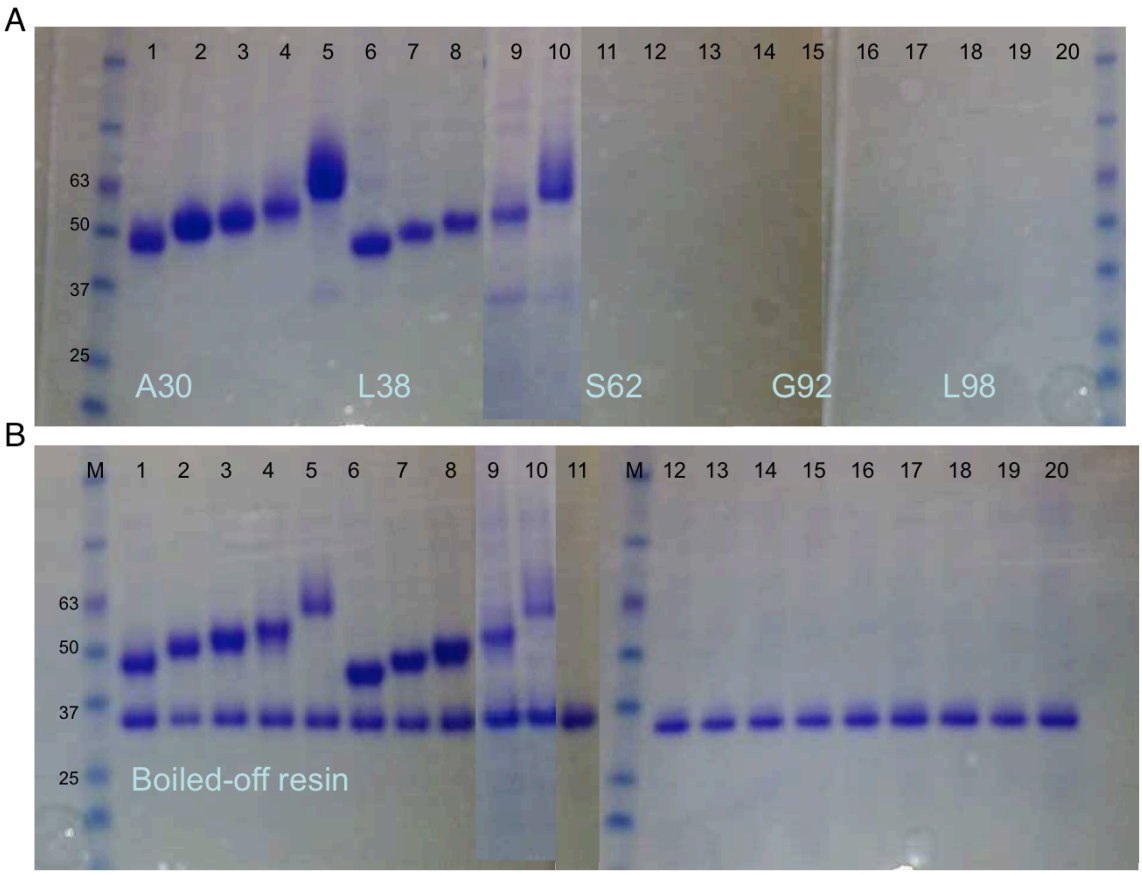


Figure 2.5 Coomassie stained gels of a series of Epa1 NT fragments (20)

Expressed in adherent HEK293 GNTI^{-} cells and purified by Ni2+-NTA affinity, using the mammalian expression plasmid pαSHP-H. (For clarity, composite figures were created to maintain continuity of results). Numbered lanes correspond to different Epa1p-Nterm domains as indicated. Two sets of samples run #1-20. A. Elutions – Imidazole (250 mM) and B. Resin fractions (boiled-off in 1x SDS-loading buffer).

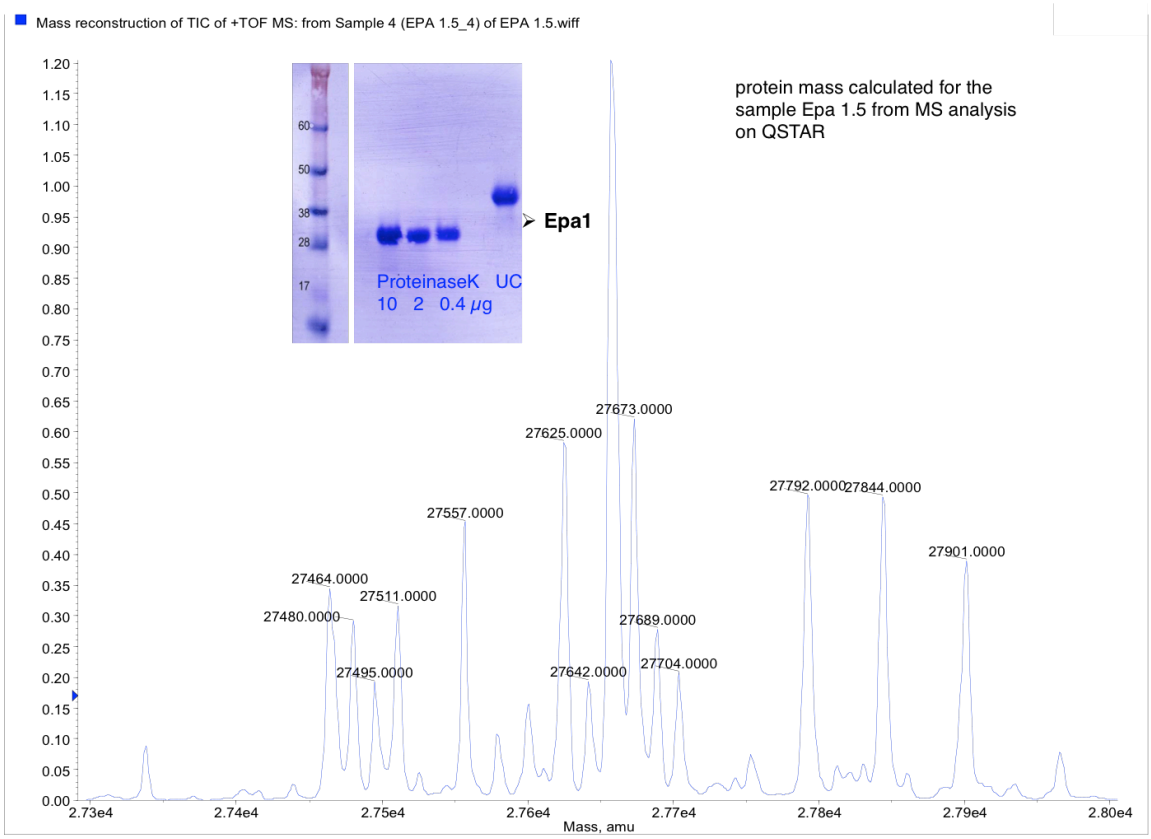


Figure 2.6 Mass Spectrometry Analysis and Coomassie stain insert of Limited Proteolysis of Epa1.

Coomassie stained PVDF blot of Epa1.5 protein treated with various concentrations of Proteinase K along with an UC (uncut) control. These bands were excised and processed for mass-spectrometry analysis. Mass spectrum analysis yields a "peptide fingerprint" that identifies a very stable Epa1 domain corresponding to **Serine 40 to Isoleucine 283**.

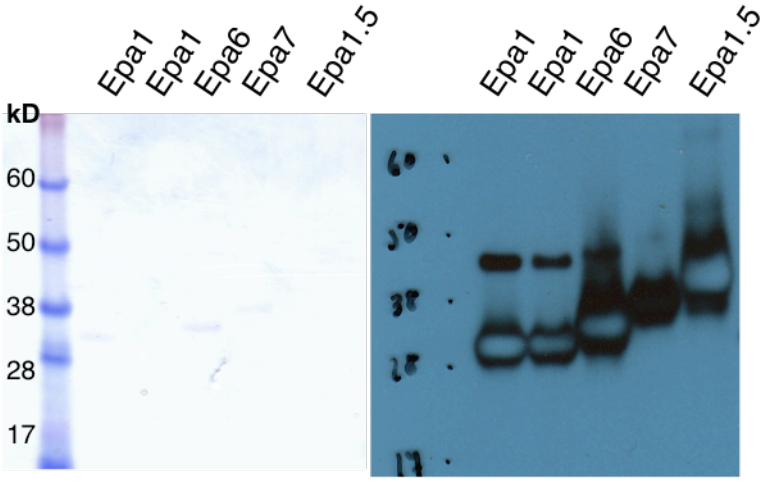


Figure 2.7 Coomassie and Western of Epa 1, 6, 7 and Epa1.5 transfection trials.

Small-scale mammalian expression of Epa proteins is low yielding. Lane descriptions are detailed in Table 2.3. Coomassie gel shows faint banding not captured in the image. Western blot indicates domain size difference between Epa1 and Epa1.5 as expected. Also evident on Western are presumably distinct glycosylation patterns between Epa1, Epa6 and Epa7 resulting in pronounced separation in PAGE gel migration patterns. Apparent 'burn thru' or halo effect of bands on film is due to overloading or overexposed Western blot.

1° ab. anti-NT Epa1 #3638 (rabbit polyclonal lab stock, dilution 1-10K).

2° ab. Mouse anti-rabbit ab. (dilution 1-10K).

Film exposure ~10sec.

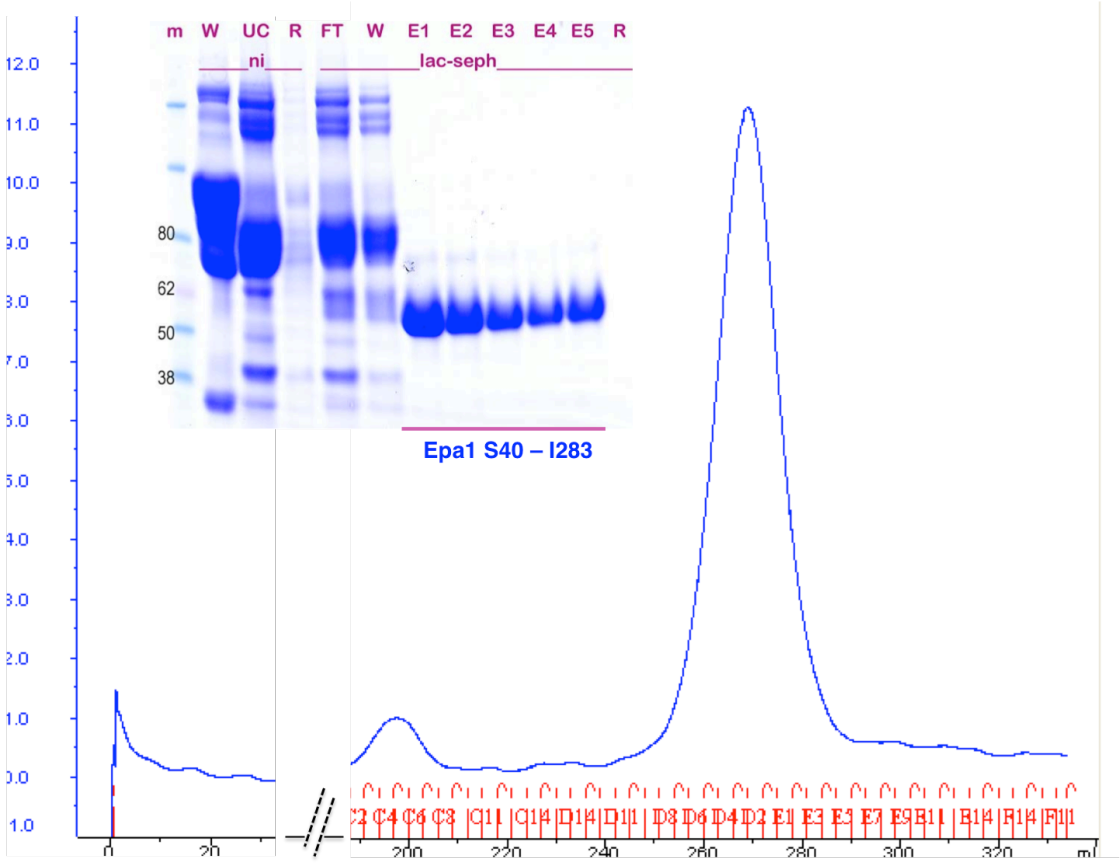


Figure 2.8 2-Step Purification and Size Exclusion Chromatography

Coomassie gel insert of **Epa1 (S40-I283)** production in HEK293 GNTI¯ cells. Sample fractions collected along IMAC (immobilized metal affinity chromatography) step followed by sample fractions of the lectin-binding step using Lactose-sepharose resin. The lac-sepharose elutions were passed over a diagnostic size-exclusion chromatography column (S75). The UV (ultraviolet-280nm) trace peak indicates appropriate sizing profile considering glycosylation features of Epa 1. Functionality is also inferred by the absence of protein in the initial void volume. (M- prestained protein molecular weights standards, W- wash, UC- uncut, R- resin, FT- flow thru, E- elutions).

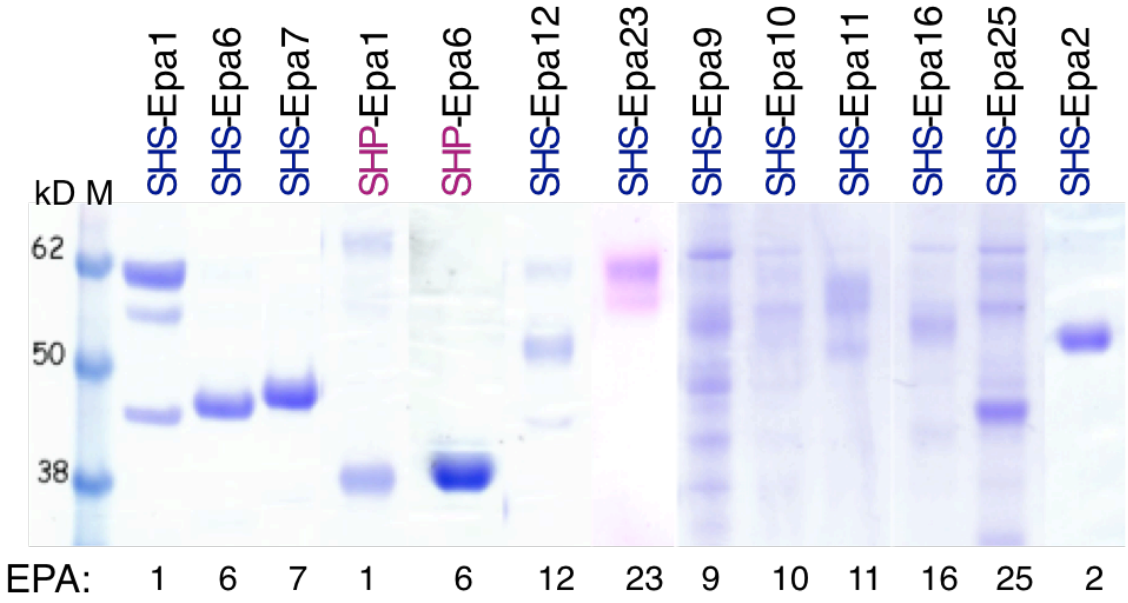


Figure 2.9 Coomassie stained gel of EPA family of NT domain fragments. (Composite figure)

Proteins were made as secreted His-Sumo1 fusions in HEK-293 cells, using the pαSHS mammalian expression plasmid. Epa1 and Epa6 were also made as non-fusions with only a His tag using pαSHP (Epa7 not shown). Not all proteins were successfully expressed, those confirmed by Western analysis continued on for glycan array analysis. (SHS- signal sequence 8xHis Sumo1, SHP- signal sequence 8xHis prescission site). Glycan-chip experiments were conducted by the CFG (Consortium for Functional Glycomics). http://www.functionalglycomics.org/fg/

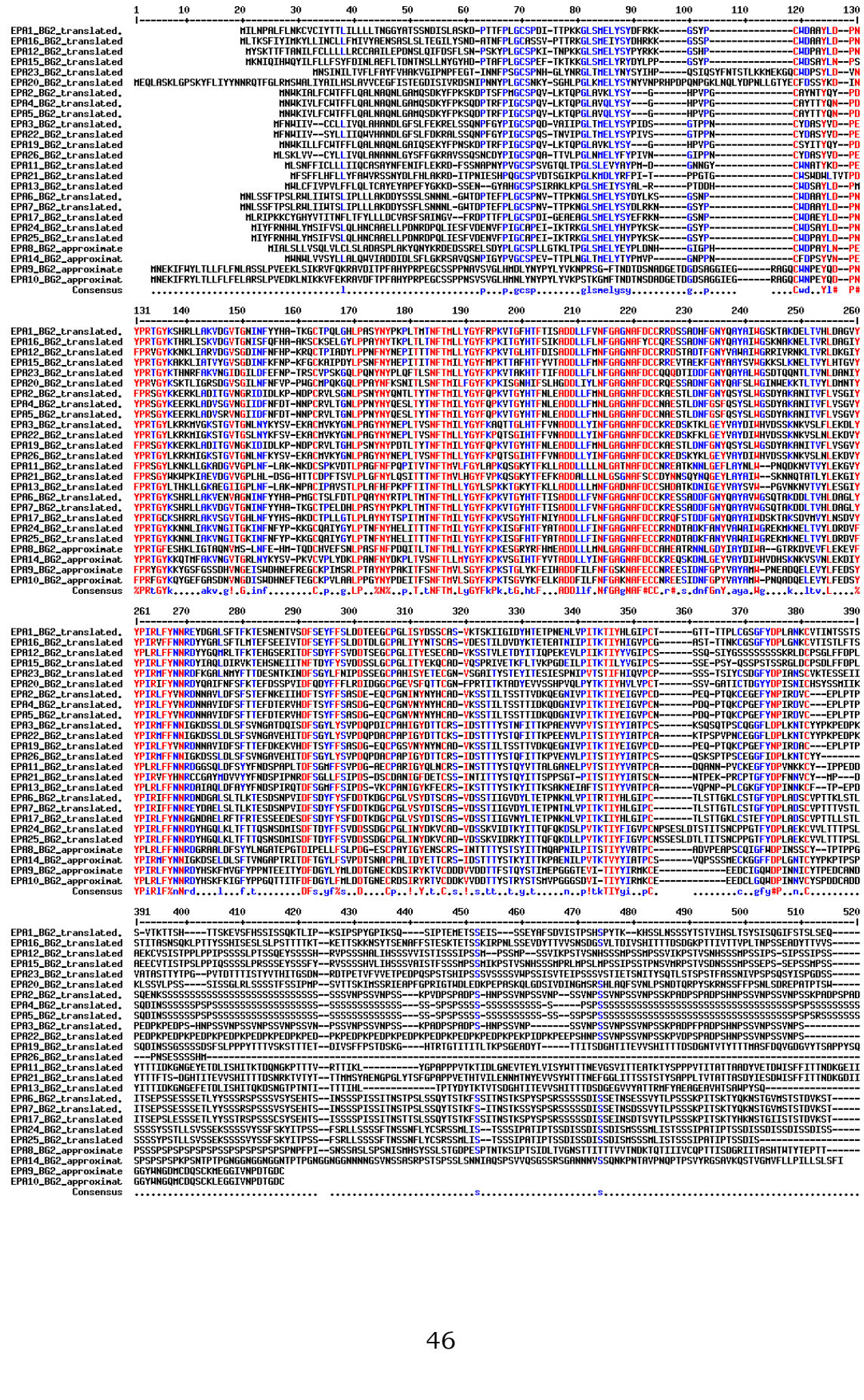


Figure 2.10 EPA Gene Family Amino Acid Sequence Alignment.

C. glabrata EPA gene family sequences – 25 in total, were assembled into a multi-sequence alignment program T-coffee. Numbered amino acids are labeled on top correspond to the longest sequence present (Epa20) and consensus amino acids are displayed below. Higher homology is depicted in red lettering and conserved residues of lesser consensus are depicted in blue lettering. Many key residues identified as part of the PA-14 domain within the N-terminal domain of Epa1 are also laid out along the consensus.

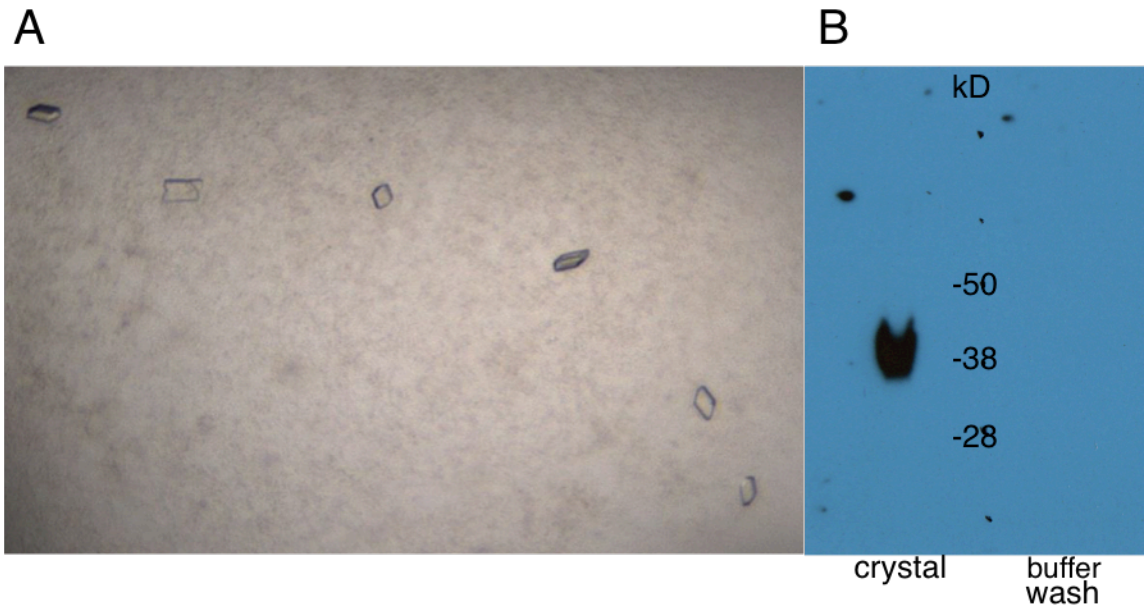
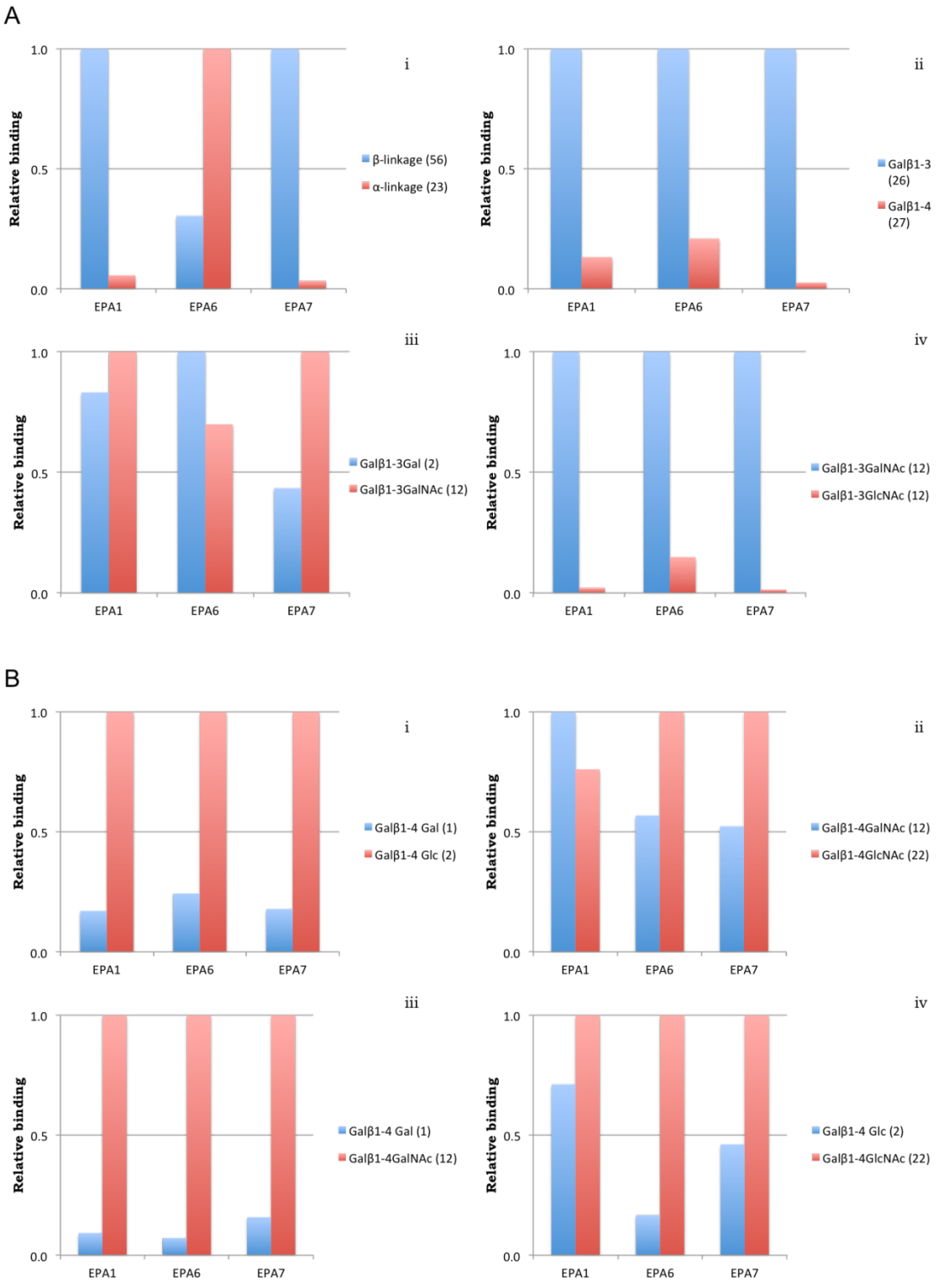
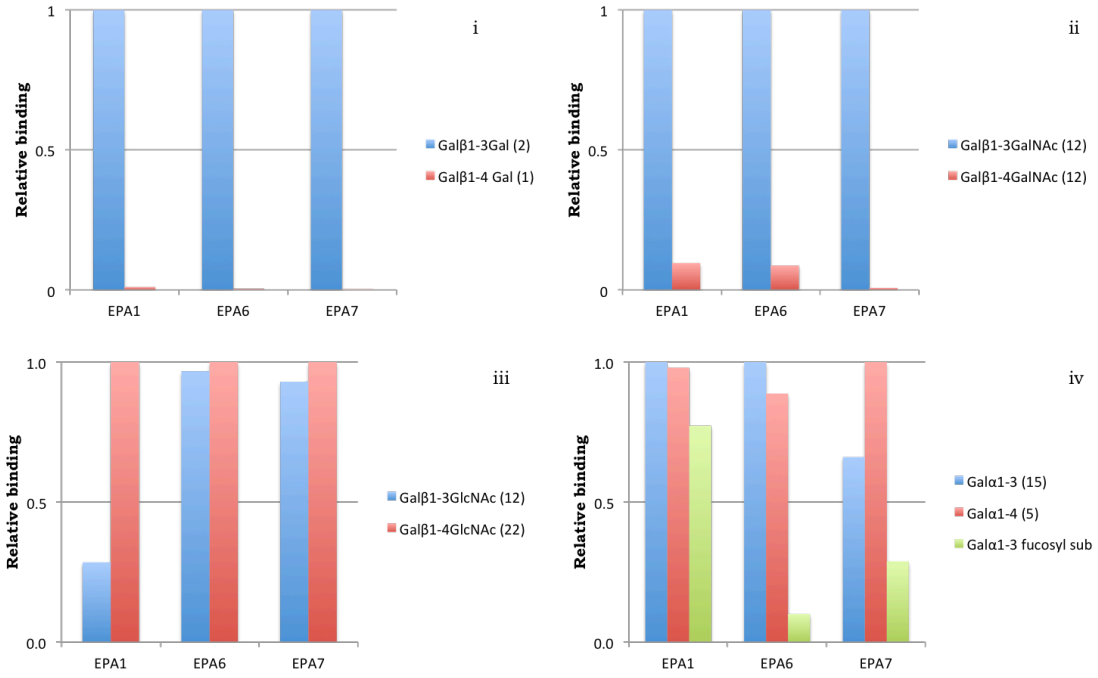


Figure 2.11 Epa7 protein crystals and Western blot analysis.

A. Hanging drop vapor diffusion technique was used to obtain visible crystals. Initial crystallization trays were set using the Qiagen JCSG+ (Joint Consortium for Structural Genomics) commercial screens. Crystals were differentiated after about one week at 30°C in drops with a protein concentration of 6 mg / ml in wells and a reservoir solution of 50% PEG 4000, 0.1M NaCl, 0.2M $Li(SO_4)_2$. B. Western of protein crystal of Epa7; crystals were recovered and run on PAGE for subsequent confirmation with anti EpaNT rabbit polyclonal ab (lab stock #3638). The lane on the right side of the protein standards marker is a negative control, the reagent buffer shows the crystal is not a consequence of salt supersaturation.


A
i
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
β-linkage (56)
α-linkage (23)
ii
Galβ1-3 (26)
Galβ1-4 (27)
iii
Galβ1-3Gal (2)
Galβ1-3GalNAc (12)
iv
Galβ1-3GalNAc (12)
Galβ1-3GlcNAc (12)
B
i
Galβ1-4 Gal (1)
Galβ1-4 Glc (2)
ii
Galβ1-4GalNAc (12)
Galβ1-4GlcNAc (22)
iii
Galβ1-4 Gal (1)
Galβ1-4GalNAc (12)
iv
Galβ1-4 Glc (2)
Galβ1-4GlcNAc (22)

C


i
Relative binding
1
0.5
0
EPA1
EPA6
EPA7
Galβ1-3Gal (2)
Galβ1-4 Gal (1)
ii
Relative binding
1
0.5
0
EPA1
EPA6
EPA7
Galβ1-3GalNAc (12)
Galβ1-4GalNAc (12)
iii
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
Galβ1-3GlcNAc (12)
Galβ1-4GlcNAc (22)
iv
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
Galα1-3 (15)
Galα1-4 (5)
Galα1-3 fucosyl sub

D

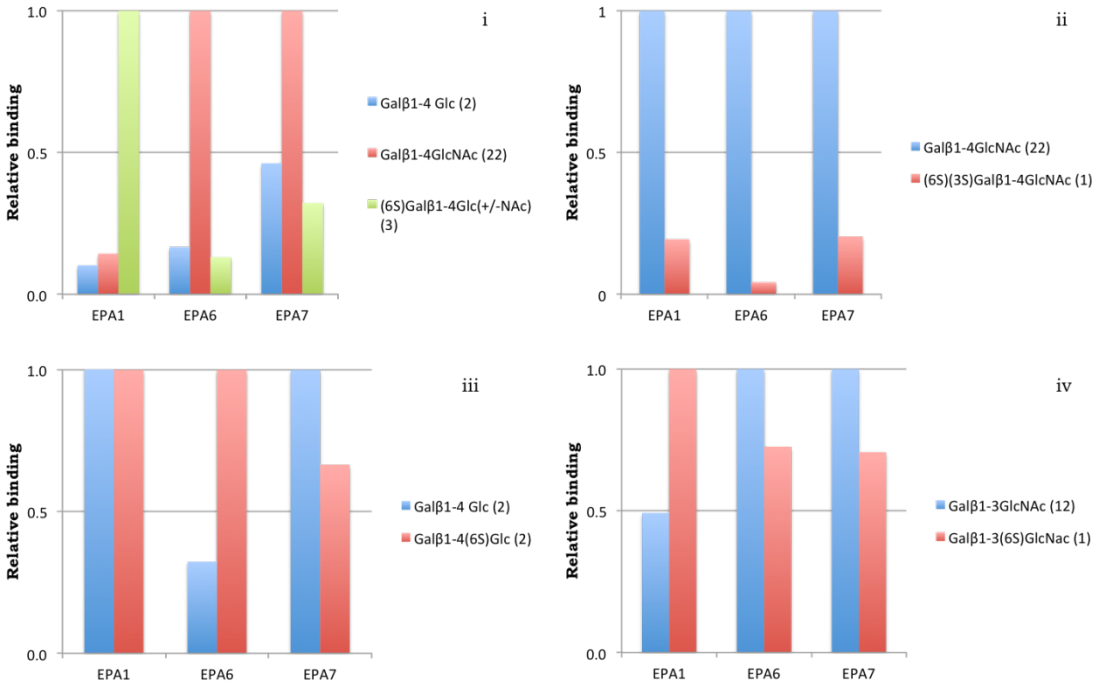
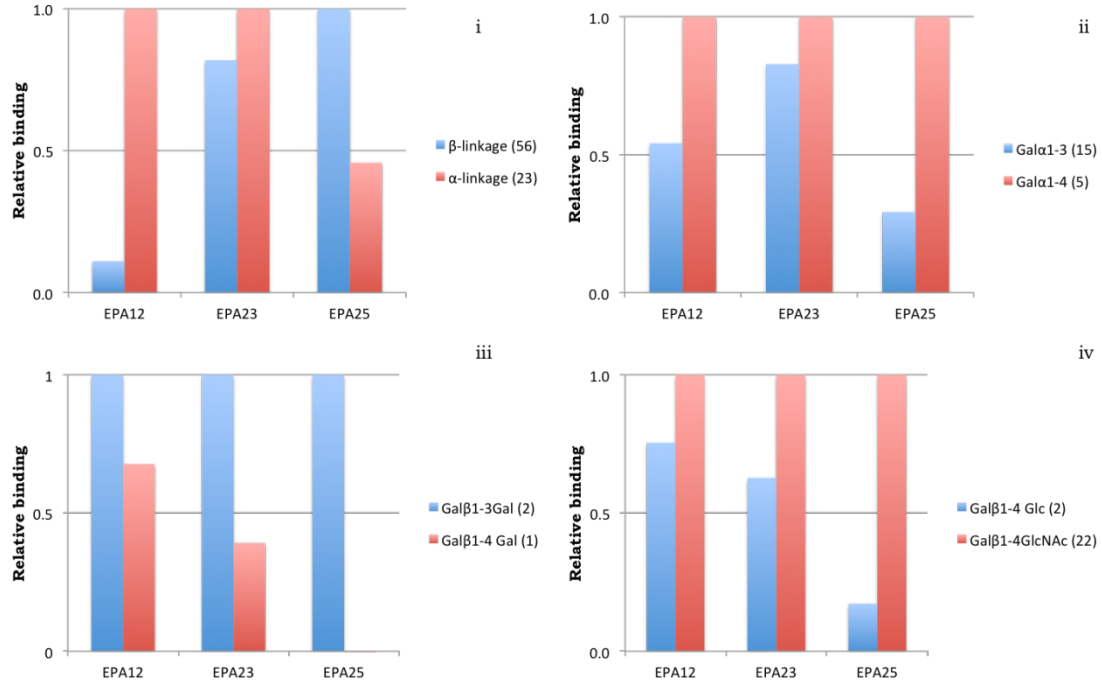

i
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
Galβ1-4 Glc (2)
Galβ1-4GlcNAc (22)
(6S)Galβ1-4Glc(+/-NAc) (3)
ii
Relative binding
1
0.5
0
EPA1
EPA6
EPA7
Galβ1-4GlcNAc (22)
(6S)(3S)Galβ1-4GlcNAc (1)
iii
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
Galβ1-4 Glc (2)
Galβ1-4(6S)Glc (2)
iv
Relative binding
1.0
0.5
0.0
EPA1
EPA6
EPA7
Galβ1-3GlcNAc (12)
Galβ1-3(6S)GlcNac (1)

Figure 2.12 Epa1, Epa6 and Epa7 comparisons of normalized averaged rfu (relative fluorescent units) for various sugar categories from glycan microarrays.

Numbers in parenthesis indicates the number of glycans used in the comparison. A. Analysis of the anomeric configuration of each linkage (α or β) between the non-reducing end (terminal galactose) and the penultimate residue on the binding specificity and of the linkage position between the carbon numbers of the hydroxyls involved. (i) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β-linkage (Galβ) or α linkage (Galα). (ii) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) or β1-4 (Galβ1-4) linkage. (iii) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) linkage and the various penultimate residues (Gal, GalNAc). (iv) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) linkage and varying penultimate residues galactose or glucose and the influence of N-acetylated (NAc) modifications (GalNAc, GlcNAc). B. Analysis of the Galβ1-4 configuration subcategory with various penultimate residues on binding specificity. (i) Comparison of glycans containing a penultimate glucose or galactose residue (Galβ1-4Gal, Galβ1-4Glc). (ii) Comparison of glycans containing a penultimate N-acetylated modification on either a glucose or galactose residue (Galβ1-4GalNAc, Galβ1-4GlcNAc). (iii-iv)

Comparison of glycans and N-acetyl modification of galactose or glucose (Gal β1-4Gal, Galβ1-4GalNAc) and (Galβ1-4Glc, Galβ1-4GlcNAc). C(i-iii). Analysis of the influence of the type of bond β1-3 or β1-4 attached to various penultimate residues (Gal, GalNAc, GlcNAc). C(iv). Analysis of the linkage position of the α conformer category. Comparison of glycans of the α configuration class and the linkage types (Galα1-3, Galα1-4, Galα1-3 fucosyl substitution) D. Analysis of the modified terminal disaccharides by sulfation at the non-reducing galactose end, and/or the penultimate positions as well as the influence of N-acetylation on the penultimate residue (6S, 3S, NAc). (i) Comparison of glycans with the presence of sulfation at the terminal galactose and N-acetylation modifications on the penultimate glucose. (ii) Comparison of glycans with a (Galβ1-4GlcNAc) and the presence of two sulfation modifications (6S, 3S) at the terminal non-reducing end. (iii, iv) Comparison of glycans with varying positions of the galactose-β linkage (Galβ1-4(6S)Glc) and sulfated penultimate residue with modified glucose (N-acetylation).

A


i
Relative binding
1.0
0.5
0.0
EPA12
EPA23
EPA25
β-linkage (56)
α-linkage (23)
ii
Galα1-3 (15)
Galα1-4 (5)
iii
1
Galβ1-3Gal (2)
Galβ1-4 Gal (1)
iv
Galβ1-4 Glc (2)
Galβ1-4GlcNAc (22)

B

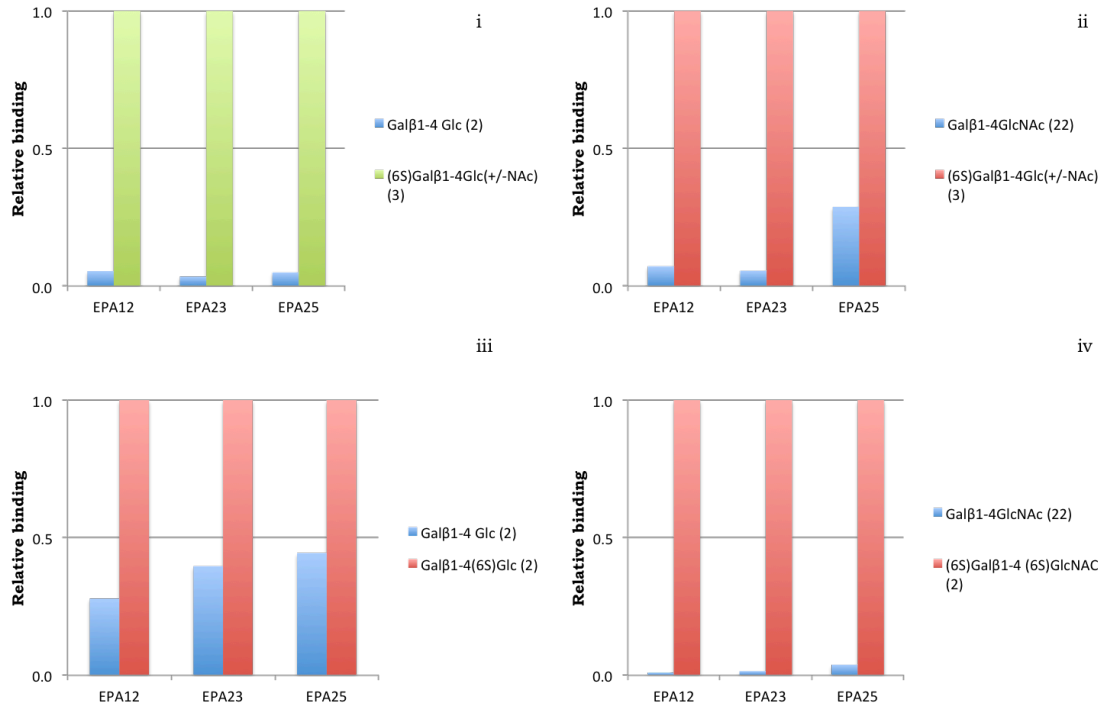

i
Relative binding
1.0
0.5
0.0
EPA12
EPA23
EPA25
Galβ1-4 Glc (2)
(6S)Galβ1-4Glc(+/-NAc) (3)
ii
Galβ1-4GlcNAc (22)
(6S)Galβ1-4Glc(+/-NAc) (3)
iii
Galβ1-4 Glc (2)
Galβ1-4(6S)Glc (2)
iv
Galβ1-4GlcNAc (22)
(6S)Galβ1-4 (6S)GlcNAC (2)

Figure 2.13 Epa12, Epa23 and Epa25 comparisons of normalized averaged rfu (relative fluorescent units) for various sugar categories from glycan microarrays.

Numbers in parenthesis indicates the number of glycans used in the comparison. A. Analysis of the anomeric configuration of each linkage (α or β) between the non-reducing end (terminal galactose) and the penultimate residue on the binding specificity and of the linkage position between the carbon numbers of the hydroxyls involved. (i) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β-linkage (Galβ) or α linkage (Galα). (ii) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) or β1-4 (Galβ1-4) linkage. (iii) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) linkage and the various penultimate residues (Gal, GalNAc). (iv) Comparison of glycans containing a terminal galactose residue linked to the penultimate residue by a β1-3 (Galβ1-3) linkage and varying penultimate residues galactose or glucose and the influence of N-acetylated (NAc) modifications (GalNAc, GlcNAc). B. Analysis of the Galβ1-4 configuration subcategory with various penultimate residues on binding specificity. (i) Comparison of glycans containing a penultimate glucose or galactose residue (Galβ1-4Gal, Galβ1-4Glc). (ii) Comparison of glycans containing a penultimate N-acetylated modification on either a glucose or galactose residue (Galβ1-4GalNAc, Galβ1-4GlcNAc). (iii-iv)

Comparison of glycans and N-acetyl modification of galactose or glucose (Galβ1-4Gal, Galβ1-4GalNAc) and (Galβ1-4Glc, Galβ1-4GlcNAc).

Construct #	N-term	C-term	DNA bp	MW Daltons
1	A30	L255	675	24750
2	A30	S274	732	26840
3	A30	E288	774	28380
4	A30	P307	831	30470
5	A30	H346	948	34760
6	L38	L255	650	23833
7	L38	S274	708	25960
8	L38	E288	750	27500
9	S62	L255	600	22000
10	S62	S274	636	23320
11	S62	E288	678	24860
12	L98	L255	471	17270
13	L98	S274	528	19360
14	L98	E288	570	20900

Table 2.1 EPA 1 – NT domain expression constructs and expected protein molecular weights (Daltons).

Epa1 lectin domain fragments were cloned into a modified pET28a plasmid as 6xHis-Sumo1 N-terminal fusions. Directional cloning of fragments made by *BamHI* – *Xho*I restriction sites. Constructs were expressed in BL-21 cells (Cloning and expression details in Materials & Methods section).

	Buffer (20mM)	pH	NaCl [mM]	AA (100mM)	Detergent (0.1%)
1	Bis-tris	6.5	150		
2	Tris-Cl	8.0	150		
3	Bis-tris	6.5	500		
4	Tris-Cl	8.0	500		
5	Bis-tris	6.8	150	Arginine	
6	Bis-tris	6.8	150	Glycine	
7	Tris-Cl	8.0	150	Glycine	
8	Bis-tris	6.5	150		NP40
9	Bis-tris	6.8	150		Triton X-100
10	Bis-tris	6.8	500		Triton X-100
11	Bis-tris	8.0	500		NP40

Table 2.2 Varying parameters for solubility.

A series of biophysical parameters were tested to improve folding behavior of Epa1p. The 11 buffers focused on: salt, pH, reducing capacity and detergent solubility.

	EPA	Expression domain	4μg DNA (μl)	PEI 2:1 (8μg)	HSFM (μl)
1	Epa1 no stop	S40__	5	8 μl	500
2	Epa1	S40 - I283	5.3	8 μl	500
3	Epa6	S40 - V283	6	8 μl	500
4	Epa7	S40 - V283	7.2	8 μl	500
5	Epa1.5	A30 - H346	3.6	8 μl	500

Table 2.3 General small-scale transfection protocol. Expression tests for Epa 1, 6, 7 and Epa1.5 performed in HEK293 GNTI- cells using 6-well plates.

Epa proteins expressed as secreted fusion proteins (pα SHS – Signal sequence, 8xHis, Sumo-1). Transfections carried out when cells are approximately 80% confluent. (PEI-polyethylinimine, HSFM- hybridoma serum free media). Protocol is detailed in methods.

- DNA and PEI (mixed at 1:2 mass ratio) complexes are allowed to proceed in HSFM, incubated for 10 min, then added back to 1xPBS washed cells.
- Cells are harvest 4da post transfection
- Ni^{2+} purification
- Ulp1 digest overnight
- Lac-sepharose purification

EPA	Domain	Sequence
EPA1	S40 - I283	ggtggatccTCAAAAGACCCCACAACCTTTC
		ctgctcgagTTAAATACCAATTATTTTAGATGTT
EPA2	S34 - Q327	cacagatctTCCAAGGATCCGACAAGTTTTCC
		ctgctcgagttaCTGACTTGGTGTTGGAAGAGGTT
EPA3	E30 - D330	cacggatccGAATTAAGTTCCCAGAACCCATT
		ctgctcgagttaGTCCTCGGGTTTTGGATAGTAG
EPA4	F31 - Q327	cacggatccTTTCCTAAATCGCAAGACCCAAC
		ctgctcgagttaTTGACTTGGTGTAGGAAGAGGT
EPA5	F31 - Q327	cacggatccTTTCCTAAATCGCAAGACCCA
		ctgctcgagttaTTGACTTGGTGTAGGAAGAGGT
EPA6	W40 - V283	cgcggatccTGGACGGACCCTACTGAATTT
		ctgctcgagttaAACACCGATGATTGTAGATGA
EPA7	W40 - V283	cgcggatccTGGACGGACCCTACTGAATTT
		ctgctcgagttaAACACCGATGATTGTAGATGA
EPA8	A24 - G329	ggtagatctGCAAAGTATCAAAATTACAAGAGA
		ctgctcgagttaGCCCGGAGGAGTAGGAGTATAA
EPA9	N363 - S28	cacggatccAGCATCAAGAGAGTGTTCCAG
		ctgctcgagttaATTAGCGCAATCTTCGGGGGT
EPA10	N28 - Y335	ggtagatctAATATTAAAAAAGTGTTCGAGAAG
		ctgctcgagttaATAATAGATGGTTATTACATCTG
EPA11	D30 - Y325	ggtagatctGATTTTTCATCAAATGCTCCTAA
		ctgctcgagttaATAATCATCCTCTGGAGGAATGT
EPA12	S31 - L349	ggtagatctTCCCTACAAATATTTGATTCC
		ctgctcgagttaGAGAGGAGGTGTTGATATACT
EPA13	D30 - I323	gacagatctGACTCCTCCGAAAACGGTTACG
		ctgctcgagttaAATGTAATCAGGTTCAGGAGTAAA
EPA	**Domain**	**Sequence**
EPA14	S331 - S30	cacggatccTCAGCGGTTCAAAGTAATCCA
		ctgctcgagttaGGATGGTGTTGGTTTTGGGTAG
EPA15	Y37 - L348	cacggatccTATGGATATCATGATCCAACGGC
		ctagtcgacttaGAGAGAAGGTGTTGATATGGTA
EPA16	S30 - T337	cacggatccAGTTTGTCCCTGACAGAAGGC
		ctgctcgagttaAGTGAATAATGTGCTTATCGTTAC
EPA17	S30 - T335	cacggatccAGTTTTTCTGCAATTAATGGCG
		ctgctcgagttaAGTCGAGAGTAATGTTGTTGGC
EPA19	Y30 - Q327	gacagatctTACTTTCCTAACTCTAAGGATCCA
		ctgctcgagttaTTGACTTGGCGTTGGAAGAGG
EPA20	T50 - I377	cacggatccACAGAAGGAGATATAAGTATTG
		ctgctcgagttaGATTATCATAGAGCTATAAGAATG
EPA21	T31 - T324	cacggatccACACCAAATATTGAGAGCCAC
		ctgctcgagttaTGTGGTATAATCTGGCATATAGC
EPA22	A30 - P331	cacggatccGCATTAAGTTCACAGAACCCA
		ctgctcgagttaTGGGTCCTCAGGTTTTGGATA
EPA23	E27 - I339	cacggatccGAGGGCACAATCAATAATTTTCC
		ctgctcgagttaTATCTCCGAAGATTCTGTCTT
EPA24	F39 - S312	cacggatccTTCGTTGACGAGAACGTGTTTC
		ctgctcgagttaTGATTCGCTTGGGTTACAGGG
EPA25	F39 - R282	cacagatctTTCGTTGACGAGAACGTGTTT
		ctgctcgagttaTCTATCAATAACTTTTGATGAATC
EPA26	S33 - T276	cacggatccTCGCAAAGCAATTGTGACTATC
		ctgctcgagttaTGTACTGTATGTTGTTGTCGAATC

Table 2.4. List of oligos for Epa family of N-term fragments expressed as His-Sumo fusions.

Above oligos made with *BamHI* and *XhoI* restriction sites for directional cohesive-end cloning, compatible with Dan Leahy's (JHU) pα SHP-H or pα SHS (Cormack-lab Sumo version) mammalian expression plasmids. Alternative *BglII* and *SalI* restriction sites were also used in some cases for overhang compatibility.

Table 2.5 Glycan Microarray Table.

Glycan array data from purified N-terminal domains of Epa1p, Epa6p, Epa7p, Epa12p, Epa23p and Epa25p expressed in HEK 293 ($GNTI^-$) cells. Data was obtained at two different protein concentrations as indicated (200 µg/ml and 50 µg/ml). Numbers indicate average relative fluorescence units (rfu) for each sample. Rfu values of the top 5% are shown in bold for Epa1, Epa6, Epa7 as well as the top 5% for Epa12, Epa23, Epa25.

Chart Number	Glycan	200 µg/ml Epa1	200 µg/ml Epa6	200 µg/ml Epa7	200 µg/ml Epa12	200 µg/ml Epa23	200 µg/ml Epa25	90 µg/ml Epa1	50 µg/ml Epa6	50 µg/ml Epa7
136	Neu5Aca2-6(Galb1-3)GalNAca-Sp8	**9731**	2515	**7012**	9	22	**234**	**557**	195	**4229**
137	Neu5Aca2-6(Galb1-3)GalNAca-Sp14	**4443**	1770	**4226**	17	16	**593**	**54**	69	**2308**
140	Galb1-3GalNAca-Sp8	**3620**	**4817**	**2408**	2	10	6	**395**	**765**	**1472**
146	Galb1-3GalNAcb1-4Galb1-4Glcb-Sp8	**1686**	2446	**14186**	**34**	18	**674**	**541**	**2064**	**3186**
141	Galb1-3GalNAca-Sp14	**1568**	**4342**	**1572**	8	9	**280**	**67**	452	**567**
142	Galb1-3GalNAca-Sp16	**1550**	2139	**888**	13	10	7	**68**	134	**503**
301	Galb1-3Galb1-4GlcNAcb-Sp8	**1380**	**5607**	**3780**	24	25	**84**	**162**	**2981**	**620**
412	Galb1-3GalNAcb1-4(Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	**1283**	**4230**	**1269**	7	11	6	2	449	**622**
135	GlcNAcb1-6(Galb1-3)GalNAca-Sp14	**1171**	2689	**987**	12	8	11	10	126	**418**
167	Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp8	**856**	**4228**	**3095**	2	15	8	**342**	**995**	**495**
42	(6S)Galb1-4Glcb-Sp0	**819**	57	32	**135**	**317**	**227**	14	49	9
138	Neu5Acb2-6(Galb1-3)GalNAca-Sp8	**810**	642	**1053**	3	9	6	1	15	**473**
175	GlcNAca1-6Galb1-4GlcNAcb-Sp8	**718**	1246	**556**	16	7	**679**	1	64	**127**
382	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAc-Sp0	**705**	**6887**	**6014**	7	11	7	**18**	**2733**	**711**
305	Galb1-4GlcNAca1-6Galb1-4GlcNAcb-Sp0	**670**	**3818**	**522**	23	27	**854**	2	498	**116**
143	Galb1-3GalNAcb-Sp8	**638**	1384	**7532**	13	16	**240**	**867**	**938**	**2103**
205	GlcAb1-6Galb-Sp8	**558**	1339	23	**242**	19	**53**	**240**	227	21
568	Galb1-3GlcNAcb1-6(Galb1-3)GalNAc-Sp14	**539**	2450	**1077**	16	17	2	3	200	**407**
43	(6S)Galb1-4Glcb-Sp8	**472**	75	28	**136**	**332**	**171**	**67**	30	13
19	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3)GalNAca-Sp8	**377**	**6312**	**672**	22	11	20	**23**	416	**140**
539	Galb1-3GalNAcb1-3Gal-Sp21	**362**	675	**14660**	13	9	**357**	**716**	**1627**	**2030**
298	(6S)Galb1-4(6S)GlcNAcb-Sp0	**357**	178	17	**1031**	**1007**	**371**	0	28	10
45	(6S)Galb1-4(6S)Glcb-Sp8	**354**	186	16	**813**	**1257**	**2067**	12	24	15

144	Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	**305**	**5042**	**4116**	7	9	12	1	**1455**	**541**
147	Galb1-3Galb-Sp8	**282**	3381	**775**	15	8	19	4	**999**	**334**
306	Galb1-4GlcNAcb1-6Galb1-4GlcNAcb-Sp0	**279**	1997	172	10	18	16	**43**	128	39
509	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)Galb1-4GlcNAcb1-4(Gal b1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAc-Sp21	**254**	**5460**	**701**	-1	11	8	8	363	**201**
134	GlcNAcb1-6(Galb1-3)GalNAca-Sp8	**251**	2325	**2433**	23	21	**25**	**41**	487	**315**
44	(6S)Galb1-4GlcNAcb-Sp8	**241**	57	13	**112**	**251**	**87**	9	19	19
159	Galb1-4GalNAcb1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	**181**	455	57	2	3	**652**	-1	25	21
132	Galb1-4GlcNAcb1-6GalNAca-Sp8	**180**	246	**286**	15	7	-1	**33**	25	45
356	(6S)GlcNAcb1-3Galb1-4GlcNAcb-Sp0	**170**	1188	228	9	15	7	**155**	74	48
573	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	**148**	796	58	11	15	15	**20**	50	45
145	Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	**139**	332	**6671**	4	12	22	10	**880**	**373**
591	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	**136**	1479	**1131**	25	17	13	**16**	90	**178**
111	Gala1-3GalNAca-Sp8	**134**	**3888**	233	**155**	**131**	**113**	**43**	**1558**	**111**
583	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	**131**	**4454**	181	13	**39**	13	7	433	49

439	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Galb1-4GlcNAcb1-4(Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	**127**	**4912**	**700**	18	9	10	**21**	343	**140**
388	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(Galb1-4GlcNAcb1-4(Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	**127**	**5021**	**349**	2	7	10	1	314	**69**
381	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4Glc-Sp0	**125**	**5393**	**3652**	2	3	11	5	**2273**	**462**
195	GlcNAcb1-6Galb1-4GlcNAcb-Sp8	**101**	110	64	8	6	21	**132**	53	18
100	GalNAcb1-4GlcNAcb-Sp8	**96**	33	57	**114**	**136**	**23**	**82**	65	40
245	Neu5Aca2-6(Neu5Aca2-3Galb1-3)GalNAca-Sp14	89	11	60	7	6	5	4	26	29
81	Fuca1-4GlcNAcb-Sp8	86	22	10	**32**	**83**	7	**112**	28	10
419	GalNAca1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb1-3GalNAc-Sp14	86	318	279	**87**	**91**	20	**46**	284	**237**
20	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3)GalNAc-Sp14	85	2418	280	8	17	4	5	48	59
171	Galb1-4GlcNAcb-Sp23	85	370	44	10	25	6	-1	24	17
14	Manb-Sp8	84	106	80	25	13	3	**15**	35	**59**
156	Galb1-4(6S)Glcb-Sp0	81	329	25	19	17	9	3	16	10
173	Galb1-4Glcb-Sp8	81	58	49	7	15	9	0	18	7
22	6S(3S)Galb1-4(6S)GlcNAcb-Sp0	79	48	12	**77**	**52**	13	1	24	28
165	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	74	545	30	7	10	3	6	25	15
110	Gala1-4(Gala1-3)Galb1-4GlcNAcb-Sp8	70	**4075**	71	**63**	24	17	9	**609**	**65**
170	Galb1-4GlcNAcb-Sp8	69	283	35	8	23	4	0	32	12
64	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	69	36	195	6	5	14	-1	14	**65**
411	Galb1-4GlcNAcb1-6(Neu5Aca2-6Galb1-3GlcNAcb1-3)Galb1-4Glc-Sp21	68	909	79	3	7	13	4	21	19

585	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	67	**4012**	263	3	**42**	22	15	**592**	58
486	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	66	3529	253	13	**32**	18	10	203	57
97	GalNAcb1-3Gala1-4Galb1-4GlcNAcb-Sp0	62	52	41	**45**	**37**	11	13	48	36
289	Neu5Aca2-3Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	62	837	**1269**	8	9	9	**21**	76	**150**
549	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	60	**5008**	**327**	8	28	8	2	**818**	57
363	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAb-Sp20	57	56	24	11	20	15	5	16	15
601	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	57	844	**526**	11	21	9	7	41	**76**
24	(3S)Galb1-4(Fuca1-3)(6S)Glc-Sp0	55	53	33	**43**	29	2	13	18	25
340	GlcNAca1-4Galb1-4GlcNAcb-Sp0	54	18	22	7	17	7	2	14	10
300	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-6(Galb1-3)GalNAca-Sp14	53	124	74	10	20	5	4	25	41

501	Fuca1-2Galb1-3(6S)GlcNAcb-Sp0	52	115	74	10	7	12	3	45	35
560	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2)Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Mana1-4GlcNAcb1-4GlcNAc-Sp24	52	1602	107	7	21	**28**	10	195	35
587	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	51	**3837**	**354**	10	**69**	16	14	**611**	45
407	Galb1-3GlcNAcb1-6Galb1-4GlcNAcb-Sp0	51	539	**620**	1	11	19	**29**	61	**90**
559	GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-2Man a1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp24	50	73	35	**39**	**44**	11	15	36	24
387	Galb1-3GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-6(Galb1-3GlcNAcb1-3)Galb1-4Glc-Sp21	49	3370	41	6	10	6	-2	67	19
543	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	46	**4102**	251	9	**37**	2	-3	492	29
123	Gala1-4Galb1-4Glcb-Sp0	46	1644	30	**37**	**51**	17	**17**	**828**	19

572	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	45	393	82	9	22	7	10	59	32
510	Galb1-3GlcNAca1-3Galb1-4GlcNAcb-Sp8	44	70	10	3	4	8	-2	21	8
430	Galb1-4GlcNAcb1-6(Fuca1-2Galb1-3GlcNAcb1-3)Galb1-4Glc-Sp21	44	221	42	5	5	7	14	12	12
578	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	44	3293	200	19	**37**	11	**17**	511	48
441	Galb1-6Galb-Sp10	44	908	37	9	15	21	3	252	11
511	Galb1-3(6S)GlcNAcb-Sp8	42	339	50	0	6	16	-2	30	10
109	Gala1-3(Fuca1-2)Galb-Sp18	41	861	41	**67**	26	3	**26**	50	48
18	GlcN(Gc)b-Sp8	41	42	16	20	25	9	4	49	31
569	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	40	**4521**	229	13	**41**	10	-1	436	**62**

383	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3GlcNAcb1-3)Galb1-4Glcb-Sp0	39	2918	87	6	17	13	11	380	20
366	Fuca1-4(Galb1-3)GlcNAcb1-2Mana1-6(Fuca1-4(Galb1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	39	42	36	7	3	4	3	35	25
343	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	37	40	37	14	10	16	8	42	32
12	Galb-Sp8	37	795	36	28	21	10	**16**	230	26
9	Neu5Aca-Sp8	36	40	27	24	21	21	10	24	26
417	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3GalNAca-Sp14	36	134	46	14	13	6	8	32	25
4	GalNAca-Sp8	36	13	5	7	17	3	2	23	12
355	Galb1-3GlcNAcb1-2Mana1-6(Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	35	2834	36	2	9	7	3	102	20
576	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	35	3340	220	18	**42**	8	9	333	35
603	Neu5Aca2-6Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	35	582	**832**	12	19	6	**20**	24	**95**
437	Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Galb1-4GlcNAcb1-4(Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	34	3595	186	7	17	9	4	139	36
62	Fuca1-2Galb1-3GalNAca-Sp8	34	23	13	4	6	3	3	9	7
17	GlcNAcb-Sp8	33	26	10	9	29	6	1	32	23
85	(3S)Galb1-4(Fuca1-3)Glcb-Sp0	33	11	24	8	8	16	3	9	11
47	(6S)GlcNAcb-Sp8	32	13	10	11	14	3	2	14	8

589	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	31	3437	**594**	19	**66**	11	13	**721**	**60**
163	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	31	718	44	0	11	2	6	26	19
164	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	31	567	35	9	29	8	4	31	28
338	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-6(Neu5Aca2-3Galb1-3)GalNAc-Sp14	31	31	36	20	14	5	**15**	30	21
31	(3S)Galb1-3GlcNAcb-Sp8	31	53	44	26	13	9	4	40	32
445	(6S)Galb1-3(6S)GlcNAc-Sp0	30	71	21	**110**	**126**	**175**	2	27	7
233	GalNAcb1-4(Neu5Aca2-3)Galb1-4GlcNAcb-Sp8	30	19	14	0	2	-1	4	13	11
46	Neu5Aca2-3(6S)Galb1-4GlcNAcb-Sp8	30	16	9	9	19	8	-1	12	8
542	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	30	55	42	**34**	**37**	**30**	8	30	25

582	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp19	30	2282	**302**	20	**58**	19	14	449	36
379	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3GalNAca-Sp14	29	152	50	22	18	17	6	49	47
3	Mana-Sp8	28	43	23	16	24	9	**18**	30	44
566	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	28	1064	**471**	6	10	20	2	**645**	25
565	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	28	23	11	9	10	16	2	10	17
370	Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-4(Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	28	1760	227	2	9	8	5	94	30
7	Fuca-Sp9	28	351	150	20	16	4	6	164	**169**
106	Gala1-3(Fuca1-2)Galb1-4GlcNAc-Sp0	28	2426	21	4	4	16	1	221	12

169	Galb1-4GlcNAcb-Sp0	27	40	44	15	18	12	**18**	20	15
384	Galb1-4(Fuca1-3)GlcNAcb1-6(Galb1-3GlcNAcb1-3)Galb1-4Glc-Sp21	27	100	41	4	10	7	2	31	9
491	Galb1-3(Fuca1-4)GlcNAcb1-6GalNAca-Sp14	27	49	21	11	3	11	10	21	24
595	GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	27	1516	**1239**	3	-3	14	11	146	**114**
278	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	26	73	20	3	6	6	-2	18	12
5	GalNAca-Sp15	26	52	33	16	20	7	6	18	19
10	Neu5Aca-Sp11	25	29	19	7	9	13	0	13	10
15	GalNAcb-Sp8	25	18	19	10	19	9	4	12	19
457	Neu5Aca2-6Galb1-4GlcNAcb1-6(Fuca1-2Galb1-3GlcNAcb1-3)Galb1-4Glc-Sp21	25	27	18	4	3	3	2	13	7
380	GalNAcb1-4GlcNAcb1-2Mana1-6(GalNAcb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	25	37	16	10	14	9	1	17	14
546	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	24	34	24	25	**44**	**24**	6	29	23
592	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3)GalNAca-Sp14	24	227	46	13	**31**	19	14	38	38
1	Gala-Sp8	24	2768	25	**243**	28	**36**	7	**885**	17
325	Galb1-3GlcNAcb1-2Mana1-6(Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	24	**3977**	25	14	9	2	0	100	16
290	Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	24	57	23	3	4	11	3	10	3

58	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	24	28	38	16	**38**	19	2	32	29
579	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	23	20	17	3	22	10	9	19	16
99	GalNAcb1-4GlcNAcb-Sp0	23	53	37	**30**	**39**	6	4	53	44
13	Glcb-Sp8	23	30	29	6	29	13	7	23	27
172	Galb1-4Glcb-Sp0	23	102	21	7	6	7	0	38	12
547	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	23	29	29	13	3	11	0	16	15
41	(6P)Mana-Sp8	23	16	18	12	17	8	-2	14	16
448	Galb1-4(Fuca1-3)GlcNAcb1-6GalNAc-Sp14	23	68	28	10	13	11	-1	25	25
63	Fuca1-2Galb1-3GalNAca-Sp14	23	53	47	9	14	9	6	10	20
157	Galb1-4(6S)Glcb-Sp8	22	168	21	**31**	**36**	**27**	1	27	8
395	Gala1-3Galb1-3GlcNAcb1-2Mana1-6(Gala1-3Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp19	22	414	17	7	14	3	2	49	18
6	Fuca-Sp8	22	19	10	11	-1	11	7	30	14
438	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	22	3457	79	14	8	3	2	70	32

378	Neu5Aca2-6Galb1-4GlcNAcb1-3GalNAc-Sp14	22	32	31	7	7	14	2	20	14
160	Galb1-4GlcNAcb1-3GalNAca-Sp8	22	241	39	6	8	7	-1	19	13
464	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	22	13	15	-1	6	1	5	15	13
561	Gala1-3Galb1-4GlcNAcb1-2Mana1-6(Gala1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp24	21	**6059**	49	13	22	**23**	5	**2332**	21
483	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	21	19	22	8	16	3	4	18	19
405	Gala1-4Galb1-4GlcNAcb1-2Mana1-6(Gala1-4Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	21	**4817**	43	12	20	14	5	**2658**	20
413	Neu5Aca2-3Galb1-3GalNAcb1-4(Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	21	26	44	2	11	8	1	12	32
352	Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	21	404	60	5	15	16	3	28	33
479	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana-Sp0	21	27	27	15	13	9	1	30	17
548	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	21	55	34	10	**40**	6	12	23	43

321	GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	20	9	5	7	13	12	3	16	6
392	Galb1-3GlcNAcb1-3GalNAca-Sp14	20	26	10	**66**	8	23	-1	12	10
534	Fuca1-4(Galb1-3)GlcNAcb1-2 Mana-Sp0	20	179	28	7	7	10	5	68	33
16	GlcNAcb-Sp0	20	12	27	8	16	8	-3	18	30
254	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	20	2	13	16	12	9	0	7	9
581	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp19	20	13	40	7	5	13	2	27	16
386	Galb1-4(Fuca1-3)GlcNAcb1-6(Fuca1-4(Fuca1-2Galb1-3)GlcNAcb1-3)Galb1-4Glc-Sp21	20	63	15	9	6	-3	-1	12	16
293	Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	20	132	16	6	8	-6	-3	16	15
21	GlcNAcb1-6(GlcNAcb1-4)(GlcNAcb1-3)GlcNAc-Sp8	19	22	26	23	14	15	2	18	9
260	Neu5Aca2-3Galb1-4GlcNAcb-Sp0	19	11	5	6	16	2	2	10	18
126	Galb1-2Galb-Sp8	19	**4722**	31	10	20	21	2	**1136**	10
314	Mana1-6Manb-Sp10	19	19	11	8	-5	2	2	17	24
221	Fuca1-2(6S)Galb1-4GlcNAcb-Sp0	19	13	14	26	**35**	16	3	13	14
38	(3S)Galb-Sp8	19	24	23	2	7	6	1	15	22
489	Galb1-3GlcNAcb1-6GalNAca-Sp14	19	43	18	0	4	9	5	25	7

478	Neu5Aca2-6Galb1-4GlcNAcb1-6(Galb1-3GlcNAcb1-3)Galb1-4Glcb-Sp21	19	56	19	7	12	15	0	17	11
168	Galb1-4GlcNAcb1-6(Galb1-3)GalNAc-Sp14	19	134	**1723**	6	14	4	12	111	**155**
148	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	19	82	23	2	8	12	1	7	9
155	Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	19	74	19	2	27	3	3	13	11
125	Gala1-6Glcb-Sp8	19	**4034**	137	**243**	1	**48**	1	**1901**	24
607	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	18	49	24	20	**30**	**25**	**23**	33	21
120	Gala1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	18	1743	87	**29**	27	16	10	**2179**	24
27	(3S)Galb1-4(6S)Glcb-Sp8	18	16	19	10	20	13	7	3	16
342	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	18	15	9	13	2	8	2	6	5
102	Gala1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	18	544	17	12	11	2	3	19	8
602	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Galb1-3)GalNAca-Sp14	18	717	**400**	5	17	9	6	26	**86**
133	Galb1-4GlcNAcb1-6GalNAc-Sp14	18	30	56	18	11	10	-4	28	14
473	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	18	34	10	16	17	15	2	18	19

396	Gala1-3Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6(Gala1-3Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp19	17	101	18	14	11	10	3	19	26
493	(3S)Galb1-3(Fuca1-4)GlcNAcb-Sp0	17	8	22	9	4	5	8	6	19
36	(3S)Galb1-4GlcNAcb-Sp0	17	14	12	7	22	7	2	12	6
552	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-3GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp25	17	271	39	11	16	10	3	36	39
332	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	17	23	23	10	14	11	1	13	23
584	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	17	44	19	7	**32**	4	4	18	14
588	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	17	51	13	5	6	16	-3	31	10

391	Fuca1-2Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	17	432	142	10	6	6	4	39	41
291	Galb1-4(Fuca1-3)(6S)GlcNAcb-Sp0	17	15	13	9	7	18	10	21	17
78	Fuca1-2Galb1-4Glcb-Sp0	17	19	15	5	10	6	-1	9	10
551	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	17	2584	**455**	8	**41**	**71**	**28**	512	33
465	Neu5Aca2-6Galb1-4GlcNAcb1-4Mana1-6(GlcNAcb1-4)(Neu5Aca2-6Galb1-4GlcNAcb1-4(Neu5Aca2-6Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	17	12	4	15	12	16	2	17	11
303	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-3)Galb1-4GlcNAc-Sp0	17	3688	**386**	5	15	6	**18**	361	30
580	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	16	3085	277	13	**59**	16	7	**586**	31
128	Galb1-3GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	16	148	19	1	12	10	4	17	8
258	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4GlcNAcb-Sp8	16	40	27	13	14	17	9	31	14
80	Fuca1-3GlcNAcb-Sp8	16	21	22	10	7	8	-3	9	9

361	Fuca1-2Galb1-3GlcNAcb1-2Mana1-6(Fuca1-2Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	16	4	8	21	10	12	5	11	16
502	Fuca1-2(6S)Galb1-3(6S)GlcNAcb-Sp0	16	35	24	11	17	**35**	2	43	19
280	Neu5Acb2-6Galb1-4GlcNAcb-Sp8	16	26	7	10	7	3	1	10	9
447	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-4(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	16	62	20	9	6	12	5	23	15
323	Neu5Aca2-8Neu5Aca2-8Neu5Acb-Sp8	16	3	15	13	4	6	3	9	8
37	(3S)Galb1-4GlcNAcb-Sp8	16	8	23	16	18	13	8	13	24
577	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	16	26	14	13	20	12	3	15	13
161	Galb1-4GlcNAcb1-3GalNAc-Sp14	16	121	22	6	6	2	-2	16	9
456	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6(Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	16	80	11	10	9	18	0	12	11
116	Gala1-3Galb1-4GlcNAcb-Sp8	15	**5557**	38	16	15	4	1	**1214**	14
487	Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	15	31	15	2	7	6	2	5	16
104	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	15	37	11	21	12	11	2	15	10

225	Neu5Aca2-3Galb1-3GalNAca-Sp14	15	18	9	6	5	9	-1	4	12
450	Fuca1-2Galb1-4GlcNAcb1-6(Fuca1-2Galb1-4GlcNAcb1-3)GalNAc-Sp14	15	31	21	16	15	7	4	17	16
304	GlcNAcb1-6(Galb1-4GlcNAcb1-3)Galb1-4GlcNAc-Sp0	15	1066	38	7	16	**24**	0	50	16
329	Neu5,9Ac2a2-3Galb1-4GlcNAcb-Sp0	15	15	5	5	7	13	-2	12	15
273	Neu5Aca2-6Galb1-4Glcb-Sp0	15	10	2	5	**37**	8	-1	7	7
107	Gala1-3(Fuca1-2)Galb1-4Glcb-Sp0	15	672	25	19	14	5	-1	47	18
117	Gala1-3Galb1-4Glcb-Sp0	15	**4685**	19	**40**	7	13	-1	505	10
211	Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	15	14	3	0	10	2	1	18	18
264	Neu5Aca2-3Galb1-4Glcb-Sp0	15	22	6	12	0	9	1	21	10
425	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-3GalNAc-Sp14	15	43	25	3	13	19	3	18	17
369	Galb1-4(Fuca1-3)GlcNAcb1-6(Fuca1-2Galb1-4GlcNAcb1-3)Galb1-4Glc-Sp21	15	91	10	8	22	9	-1	15	7
152	Galb1-4(Fuca1-3)GlcNAcb-Sp0	15	119	15	8	9	10	2	18	12
517	(4S)GalNAcb-Sp10	15	8	13	5	14	7	2	16	5
316	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Mana-Sp9	15	13	18	6	14	7	1	15	6
313	MurNAcb1-4GlcNAcb-Sp10	15	9	16	6	7	8	1	17	16
57	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Man-a1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	15	20	16	12	10	2	-1	4	14
335	GalNAcb1-3Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	15	25	14	13	14	7	3	15	6
115	Gala1-3Galb1-3GlcNAcb-Sp0	15	**6180**	48	**46**	5	12	2	**843**	19

393	GalNAcb1-4(Neu5Aca2-3)Galb1-4GlcNAcb1-3GalNAca-Sp14	15	16	6	6	14	6	-1	21	15
285	Neu5Gca2-3Galb1-4Glcb-Sp0	15	23	13	8	12	8	3	16	9
162	Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	15	628	53	7	7	9	-2	36	17
95	GalNAcb1-3GalNAca-Sp8	14	7	4	2	3	4	2	18	7
307	GalNAcb1-3Galb-Sp8	14	5	6	4	15	11	1	3	15
243	Neu5Aca2-3Galb1-3(6S)GalNAca-Sp8	14	13	15	4	5	9	2	11	11
350	Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	14	411	35	8	**33**	14	0	43	14
477	Galb1-3GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	14	**12616**	41	10	6	11	2	**835**	20
429	Galb1-3GlcNAcb1-6(Galb1-3GlcNAcb1-2)Mana1-6(Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	14	1146	25	3	20	**33**	4	57	13
214	Mana1-6(Mana1-3)Mana-Sp9	14	11	9	6	15	15	1	16	8
215	Mana1-2Mana1-2Mana1-6(Mana1-3)Mana-Sp9	14	14	8	4	10	1	1	9	14
194	GlcNAcb1-6GalNAca-Sp14	14	7	11	8	3	4	5	15	19
224	Neu5Aca2-3Galb1-3GalNAca-Sp8	14	23	19	11	4	11	4	16	12
23	6S(3S)Galb1-4GlcNAcb-Sp0	14	20	15	16	27	16	6	22	24

586	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAb1-2)Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	14	41	18	15	23	7	5	14	17
35	(3S)Galb1-4(6S)GlcNAcb-Sp8	14	21	19	18	28	22	9	18	14
453	Neu5Aca2-8Neu5Aca2-3Galb1-3GalNAcb1-4(Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	14	131	31	11	21	16	8	33	18
504	GalNAcb1-4(Fuca1-3)(6S)GlcNAcb-Sp8	14	27	8	15	11	0	-2	17	10
385	Galb1-4GlcNAcb1-6(Fuca1-4(Fuca1-2Galb1-3)GlcNAcb1-3)Galb1-4Glc-Sp21	14	80	36	12	17	2	1	14	18
79	Fuca1-2Galb-Sp8	14	20	24	8	19	-4	1	18	18
538	Gala1-3(Fuca1-2)Galb1-3GalNAcb1-3Gala1-4Galb1-4Glc-Sp21	14	97	7	6	5	5	1	5	9
339	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	14	8	15	2	16	4	0	9	4
515	GalNAcb1-4(6S)GlcNAc-Sp8	14	18	21	3	7	7	4	8	12
471	Glca1-4Glca1-4Glca1-4Glcb-Sp10	14	12	9	6	6	15	2	13	4
247	Neu5Aca2-3Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	14	10	11	6	8	5	0	21	10
474	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6(Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp19	14	36	35	14	19	9	-2	26	12

50	Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	13	12	17	10	15	4	4	5	6
11	Neu5Acb-Sp8	13	17	5	17	26	**44**	-1	26	9
600	Galb1-4GlcNAcb1-3Galb1-3GalNAca-Sp14	13	50	25	6	25	13	3	11	2
98	GalNAcb1-4(Fuca1-3)GlcNAcb-Sp0	13	68	25	**31**	**33**	14	**17**	40	30
410	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)Glcb-Sp21	13	18	12	4	7	4	5	2	14
554	Neu5Aca2-8Neu5Gca2-3Galb1-4GlcNAc-Sp0	13	9	2	2	5	0	1	11	8
67	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp10	13	9	9	4	10	9	-1	11	9
113	Gala1-3GalNAcb-Sp8	13	**5448**	33	**62**	4	**25**	0	**1022**	23
55	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	13	13	24	6	5	13	-2	17	2
590	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	13	86	41	5	21	8	1	31	12
344	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	13	14	32	7	9	8	0	11	8
475	Neu5Aca2-3Galb1-3GlcNAcb1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2)Mana1-6(Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	13	28	14	2	11	11	4	22	16
8	Rhaa-Sp8	13	28	15	11	**41**	6	0	27	28
376	Fuca1-4(Fuca1-2Galb1-3)GlcNAcb1-2Mana1-3(Fuca1-4(Fuca1-2Galb1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	13	36	22	10	22	14	-1	16	13
54	Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	13	1253	64	2	7	7	3	38	21

420	Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	13	25	10	6	18	1	4	17	9
60	Fuca1-2Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp9	13	16	15	6	1	7	2	20	7
270	Neu5Aca2-6Galb1-4GlcNAcb-Sp8	13	25	9	8	4	4	2	17	5
541	GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	13	13	6	17	14	10	2	15	16
207	KDNa2-3Galb1-4GlcNAcb-Sp0	13	14	10	6	11	5	5	6	12
204	GlcAb1-3Galb-Sp8	12	24	10	6	8	10	0	7	15
228	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	12	22	6	6	2	3	0	11	18
320	Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	12	43	27	4	12	13	-1	3	9
93	GalNAca1-3Galb-Sp8	12	40	8	12	16	10	-1	16	8
70	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	12	20	9	7	15	3	0	10	4
232	GalNAcb1-4(Neu5Aca2-3)Galb1-4GlcNAcb-Sp0	12	12	4	8	4	7	1	8	6
422	GlcNAcb1-2(GlcNAcb1-6)Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	12	8	7	5	18	13	3	14	9
181	GlcNAcb1-3GalNAca-Sp14	12	14	13	11	3	4	-1	11	7
494	Galb1-4(Fuca1-3)GlcNAcb1-6(Neu5Aca2-6(Neu5Aca2-3Galb1-3)GlcNAcb1-3)Galb1-4Glc-Sp21	12	29	5	8	9	12	3	21	8
59	Fuca1-2Galb1-3GalNAcb1-3Gala-Sp9	12	24	22	9	7	9	5	24	14

449	Galb1-4GlcNAcb1-2Mana-Sp0	12	30	15	8	18	18	5	18	18
150	Galb1-3GlcNAcb-Sp0	12	80	23	14	24	12	-1	9	5
29	(3S)Galb1-3GalNAca-Sp8	12	19	9	9	9	6	2	24	1
230	Neu5Aca2-8Neu5Aca2-8Neu5Aca-Sp8	12	7	11	9	8	12	5	4	8
476	GlcNAcb1-6(GlcNAcb1-2)Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	12	22	17	19	23	-2	1	23	22
71	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	12	18	9	5	15	6	2	13	22
435	GlcNAcb1-6(GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(GlcNAcb1-4(GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	12	10	7	3	17	6	0	3	12
191	GlcNAcb1-4GlcNAcb1-4GlcNAcb1-4GlcNAcb1-4GlcNAcb1-Sp8	12	13	23	15	5	9	0	14	2
308	GlcAb1-3GlcNAcb-Sp8	12	9	9	10	7	12	0	3	5
353	GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	12	85	15	5	27	4	1	29	20
72	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb-Sp0	12	27	13	22	14	22	7	4	7
436	Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	12	309	15	14	1	10	2	16	11
337	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	12	14	18	6	12	17	1	10	11
402	Galb1-4(Fuca1-3)GlcNAcb1-3GalNAca-Sp14	12	36	12	5	4	3	1	15	27

362	Fuca1-2Galb1-4GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	12	8	23	9	10	13	0	16	0
414	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-3GalNAca-Sp14	12	261	14	6	19	1	1	25	13
354	Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	12	221	28	10	19	9	7	23	20
284	Neu5Gca2-3Galb1-4GlcNAcb-Sp0	11	10	6	7	7	11	2	9	12
25	(3S)Galb1-4Glcb-Sp8	11	21	8	11	12	7	2	4	5
229	GalNAcb1-4(Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	11	12	12	6	10	8	1	7	8
521	Neu5Aca2-6Galb1-4GlcNAcb1-2Man-Sp0	11	11	12	12	4	18	2	10	5
597	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3)GalNAca-Sp14	11	25	21	14	8	17	7	24	13
526	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-6GalNAc-Sp14	11	51	7	8	4	4	1	11	12
535	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana-Sp0	11	19	19	11	12	1	2	14	9
373	Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	11	480	12	14	12	1	4	21	12
360	KDNa2-3Galb1-3GalNAca-Sp14	11	9	14	5	3	9	0	10	12
237	Neu5Aca2-3GalNAca-Sp8	11	7	5	7	12	7	2	22	15
119	Gala1-3Galb-Sp8	11	3303	10	**230**	6	2	1	389	8
76	Fuca1-2Galb1-4GlcNAcb-Sp0	11	20	15	0	12	8	0	8	6

562	GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-3)GalNAca-Sp14	11	10	2	9	4	7	1	10	16
124	Gala1-4GlcNAcb-Sp8	11	**5034**	54	**573**	8	**237**	5	**1641**	36
69	Fuca1-2Galb1-3GlcNAcb-Sp8	11	21	23	7	16	11	0	17	15
33	(3S)Galb1-4(Fuca1-3)GlcNAc-Sp8	11	170	12	9	14	7	2	14	5
500	Fuca1-2Galb1-4GlcNAcb1-2Mana-Sp0	11	11	7	4	6	11	-2	15	13
282	Neu5Gca2-3Galb1-3GlcNAcb-Sp0	11	10	10	7	1	3	-1	9	4
34	(3S)Galb1-4(6S)GlcNAcb-Sp0	11	19	6	9	19	13	3	17	17
488	Neu5Aca2-6Galb1-4GlcNAcb1-6(Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3)Galb1-4Glc-Sp21	11	22	20	8	4	15	-3	8	7
564	GalNAcb1-4GlcNAcb1-3GalNAcb1-4GlcNAcb-Sp0	11	5	15	-1	3	7	2	10	3
198	Glca1-6Glca1-6Glcb-Sp8	11	14	9	8	-1	11	0	16	17
82	Fucb1-3GlcNAcb-Sp8	11	6	11	12	10	9	-1	14	23
68	Fuca1-2Galb1-3GlcNAcb-Sp0	11	11	18	3	9	12	-6	6	6
28	(3S)Galb1-3(Fuca1-4)GlcNAcb-Sp8	11	9	9	20	11	15	6	13	18
315	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb-Sp10	11	9	18	15	12	10	9	2	11
472	Neu5Aca2-3Galb1-4GlcNAcb1-6(Neu5Aca2-3Galb1-4GlcNAcb1-3)GalNAca-Sp14	11	16	12	3	14	4	1	7	38
178	GlcNAcb1-6(GlcNAcb1-3)GalNAca-Sp14	11	24	14	7	10	-2	4	9	14
88	GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	11	17	11	6	14	9	-3	16	19
575	GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	11	45	16	4	15	8	4	17	13

398	GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	11	293	30	3	7	2	2	15	7
86	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	10	4	14	3	13	-1	-1	6	9
200	Glcb1-6Glcb-Sp8	10	25	22	5	13	6	-2	19	8
497	Galb1-4(Fuca1-3)GlcNAcb1-2Mana-Sp0	10	28	24	11	11	9	2	17	17
514	(3S)GalNAcb1-4(3S)GlcNAc-Sp8	10	14	8	6	10	0	4	15	11
158	Galb1-4GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	10	92	22	9	16	8	3	8	8
127	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	10	87	17	10	7	10	0	23	9
309	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	20	10	7	7	4	-1	7	9
84	GalNAca1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	10	26	25	19	8	3	9	22	29
94	GalNAca1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	10	7	9	2	0	6	0	4	3
490	Gala1-3Galb1-3GlcNAcb1-6GalNAca-Sp14	10	**5177**	13	**263**	11	5	2	514	11
299	(6P)Glcb-Sp10	10	20	16	13	7	20	1	11	5
261	Neu5Aca2-3Galb1-4GlcNAcb-Sp8	10	17	9	4	12	9	4	11	9
609	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	25	25	14	**31**	7	6	17	17
244	Neu5Aca2-6(Neu5Aca2-3Galb1-3)GalNAca-Sp8	10	9	11	12	2	9	2	13	10
349	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	10	11	7	4	7	3	1	7	2

443	GalNAcb1-6GalNAcb-Sp8	10	11	13	11	11	3	3	13	5
52	GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	9	9	5	13	13	1	9	20
372	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6(Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	10	1037	8	12	7	11	4	38	15
444	(6S)Galb1-3GlcNAcb-Sp0	10	32	14	**50**	21	4	6	21	3
101	Gala1-2Galb-Sp8	10	43	21	16	2	9	1	6	13
389	GlcNAcb1-2Mana1-6(GlcNAcb1-4(GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	10	15	10	6	10	6	0	23	42
206	KDNa2-3Galb1-3GlcNAcb-Sp0	10	15	10	5	5	1	-4	6	5
553	Neu5Gca2-8Neu5Gca2-3Galb1-4GlcNAc-Sp0	10	1	8	5	7	10	2	4	9
154	Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	10	160	23	8	13	8	6	30	7
252	Neu5Aca2-3Galb1-4(6S)GlcNAcb-Sp8	10	20	7	6	11	16	-2	12	8
190	GlcNAcb1-4GlcNAcb1-4GlcNAcb1-4GlcNAcb1-4GlcNAcb1-Sp8	10	19	5	5	10	8	0	15	7
327	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	22	14	2	8	7	1	9	9
189	GlcNAcb1-4Galb1-4GlcNAcb-Sp8	10	15	11	11	6	7	3	6	6
418	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb1-3GalNAc-Sp14	10	42	18	4	5	8	-1	25	11
39	(6S)(4S)Galb1-4GlcNAcb-Sp0	10	12	7	8	18	3	4	18	10
265	Neu5Aca2-3Galb1-4Glcb-Sp8	10	58	16	17	14	7	**22**	19	9

458	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6(GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	10	9	24	0	2	7	0	4	8
129	Galb1-3(Fuca1-4)GlcNAc-Sp0	10	40	15	12	0	3	-2	13	7
506	Fuca1-2Galb1-3GlcNAcb1-6(Fuca1-2Galb1-3GlcNAcb1-3)GalNAca-Sp14	9	25	24	11	1	21	5	17	21
484	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp24	9	19	9	8	19	9	6	25	16
492	Neu5Aca2-3Galb1-3GlcNAcb1-6GalNAca-Sp14	9	22	24	8	5	15	0	14	4
292	Galb1-4(Fuca1-3)(6S)Glcb-Sp0	9	11	14	8	13	12	9	5	4
242	Neu5Aca2-3Galb1-4(Neu5Aca2-3Galb1-3)GlcNAcb-Sp8	9	11	14	7	16	4	4	21	8
611	Galb1-3GalNAcb1-4(Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp21	9	19	24	18	10	21	-2	10	15
345	GlcNAca1-4Galb1-3GalNAc-Sp14	9	20	14	-1	11	6	4	13	11
324	Neu5Gcb2-6Galb1-4GlcNAc-Sp8	9	32	10	11	14	13	1	18	11
296	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	9	15	12	9	6	9	8	3	12
322	Neu5Aca2-8Neu5Acb-Sp17	9	18	5	5	6	3	1	6	10
428	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6(Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	9	253	8	8	6	9	3	16	8
223	Fuca1-2(6S)Galb1-4(6S)Glcb-Sp0	9	18	14	5	12	20	-1	21	14
2	Glca-Sp8	9	69	40	16	24	5	3	31	20

520	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-6GalNAc-Sp14	9	6	17	10	12	1	1	14	5
246	Neu5Aca2-3Galb-Sp8	9	19	10	14	13	4	2	13	19
151	Galb1-3GlcNAcb-Sp8	9	408	33	14	4	18	4	39	23
463	Neu5Aca2-3Galb1-4GlcNAcb1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Neu5Aca2-3Galb1-4GlcNAcb1-4(Neu5Aca2-3Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	9	24	4	4	10	8	1	7	12
571	(3S)GlcAb1-3Galb1-4GlcNAcb1-2Mana-Sp0	9	18	9	5	18	15	1	9	4
48	Neu5,9Ac$_2$a-Sp8	9	19	6	16	11	16	-1	8	6
176	GlcNAcb1-2Galb1-3GalNAca-Sp8	9	23	10	7	21	7	0	9	14
524	GalNAca1-3(Fuca1-2)Galb1-4 GlcNAcb1-2Mana-Sp0	9	4	15	7	1	3	2	13	16
365	Galb1-4GlcNAcb1-2Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	9	39	17	8	22	3	1	2	14
404	Gala1-4Galb1-3GlcNAcb1-2Mana1-6(Gala1-4Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	9	**7847**	66	10	25	14	4	**3705**	20
112	Gala1-3GalNAca-Sp16	9	3432	13	0	5	18	-1	**864**	9
367	Neu5Aca2-6GlcNAcb1-4GlcNAc-Sp21	9	13	7	2	6	3	-2	10	8
440	Galb1-4Galb-Sp10	9	20	6	13	6	-1	0	19	24
74	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	9	8	10	25	9	2	3	13	11
121	Gala1-4Galb1-4GlcNAcb-Sp0	9	**4575**	56	17	14	15	3	**2206**	15
275	Neu5Aca2-6Galb-Sp8	9	11	5	11	6	12	-2	11	13
203	GlcAb-Sp8	9	24	17	5	11	6	4	11	28
406	Gala1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	9	**5195**	18	10	8	13	-1	**737**	4

415	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3GalNAca-Sp14	9	18	8	10	-1	9	0	5	14
283	Neu5Gca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	8	15	19	9	11	4	-1	10	11
433	GlcNAcb1-2Mana1-6(GlcNAcb1-4)(GlcNAcb1-4(GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	8	13	12	6	4	8	-3	3	6
610	GlcNAcb1-3Fuca-Sp21	8	8	5	5	11	10	-2	16	7
604	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	8	30	25	7	13	2	3	18	15
226	GalNAcb1-4(Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	8	16	14	2	3	10	0	15	6
545	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	8	38	15	3	12	9	6	17	18
394	GalNAca1-3(Fuca1-2)Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	8	24	10	3	10	10	2	11	11
375	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6(Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	8	1771	16	7	10	7	1	29	5
220	(3S)Galb1-4(Fuca1-3)(6S)GlcNAcb-Sp8	8	40	17	18	16	7	4	17	8

446	Fuca1-2Galb1-4 GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4GlcNAcb1-2(Fuca1-2Galb1-4GlcNAcb1-4)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	8	9	10	**32**	16	9	2	7	5
468	Gala1-3(Fuca1-2)Galb1-3GalNAca-Sp8	8	193	7	10	13	8	3	10	6
268	Neu5Aca2-6Galb1-4(6S)GlcNAcb-Sp8	8	12	7	12	11	**70**	3	15	11
253	Neu5Aca2-3Galb1-4(Fuca1-3)(6S)GlcNAcb-Sp8	8	2	14	10	6	7	-1	6	11
287	Neu5Gca2-6Galb1-4GlcNAcb-Sp0	8	25	10	0	2	10	3	12	10
461	Neu5Aca2-3Galb1-4GlcNAcb1-4Mana1-6(GlcNAcb1-4)(Neu5Aca2-3Galb1-4GlcNAcb1-4(Neu5Aca2-3Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	8	11	4	14	6	2	0	5	9
222	Fuca1-2Galb1-4(6S)GlcNAcb-Sp8	8	26	25	1	9	17	-2	12	9
249	Fuca1-2(6S)Galb1-4Glcb-Sp0	8	13	5	2	12	10	-1	17	12
49	Neu5,9Ac2a2-6Galb1-4GlcNAcb-Sp8	8	11	10	5	7	10	-2	12	13
400	Neu5Aca2-3Galb1-3GlcNAcb1-3GalNAca-Sp14	8	10	9	4	5	5	0	20	9
209	Mana1-2Mana1-6(Mana1-2Mana1-3)Mana-Sp9	8	15	7	11	3	-4	1	18	14
210	Mana1-2Mana1-3Mana-Sp9	8	22	17	13	11	10	-1	16	19
357	KDNa2-3Galb1-4(Fuca1-3)GlcNAc-Sp0	8	20	16	13	13	4	9	13	21
130	Galb1-3(Fuca1-4)GlcNAc-Sp8	8	94	24	14	10	4	0	28	18
83	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	8	11	10	11	10	8	2	17	25
269	Neu5Aca2-6Galb1-4GlcNAcb-Sp0	8	5	20	14	13	7	-1	16	9
122	Gala1-4Galb1-4GlcNAcb-Sp8	8	363	71	4	7	**31**	4	**2083**	36
149	Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp10	8	113	18	6	12	4	3	12	11

294	Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	8	126	24	8	23	2	4	16	11
248	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	8	5	10	18	5	16	2	-1	17
518	Galb1-4(6P)GlcNAcb-Sp0	8	8	14	1	6	6	2	11	9
197	Glca1-4Glca-Sp8	8	18	8	11	3	10	6	18	16
427	Fuca1-2Galb1-3GlcNAcb1-2Mana1-6(Fuca1-2Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	8	**96**	15	13	23	13	-1	28	12
480	Neu5Aca2-3Galb1-4GlcNAcb1-6GalNAca-Sp14	8	40	16	3	7	3	1	8	5
346	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	8	14	9	0	5	4	8	20	4
540	GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	10	7	9	8	6	2	7	6
61	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb-Sp8	7	11	15	9	5	6	2	9	23
250	Neu5Aca2-3Galb1-3GlcNAcb-Sp0	7	25	16	4	18	8	6	15	18
30	(3S)Galb1-3GlcNAcb-Sp0	7	18	9	11	4	8	3	8	16
187	GlcNAcb1-4-MDPLys	7	13	9	11	7	12	1	19	5
251	Neu5Aca2-3Galb1-3GlcNAcb-Sp8	7	5	9	8	15	5	5	9	8
26	(3S)Galb1-4(6S)Glcb-Sp0	7	14	16	**28**	8	-1	-2	5	12
451	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-6(Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-3)GalNAc-Sp14	7	1336	13	11	8	12	2	52	22
196	Glca1-4Glcb-Sp8	7	11	5	7	5	13	0	14	6

467	Neu5Aca2-6Galb1-4GlcNAcb1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Neu5Aca2-6Galb1-4GlcNAcb1-4(Neu5Aca2-6Galb1-4GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	7	11	3	12	6	13	0	8	5
503	Neu5Aca2-6GalNAcb1-4(6S)GlcNAcb-Sp8	7	15	27	4	10	12	0	11	21
218	Manb1-4GlcNAcb-Sp0	7	18	16	13	16	**23**	5	11	22
182	GlcNAcb1-3Galb-Sp8	7	67	9	10	8	0	0	5	18
92	GalNAca1-3GalNAcb-Sp8	7	8	29	10	4	5	7	15	9
368	Neu5Aca2-6GlcNAcb1-4GlcNAcb1-4GlcNAc-Sp21	7	4	9	4	8	8	0	11	16
528	Gala1-3Galb1-3GlcNAcb1-2Mana-Sp0	7	**5977**	21	**29**	7	4	1	**881**	12
330	Neu5,9Ac2a2-3Galb1-3GlcNAcb-Sp0	7	9	8	4	12	**34**	3	17	14
317	Mana1-2Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Mana-Sp9	7	13	11	2	9	4	1	11	5
234	GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	7	12	11	13	2	23	5	20	14
311	GlcNAcb1-4GlcNAcb-Sp10	7	4	12	10	5	17	3	14	9
498	Fuca1-2(6S)Galb1-3GlcNAcb-Sp0	7	17	17	14	3	18	0	9	22
462	Neu5Aca2-3Galb1-4GlcNAcb1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	7	10	8	4	7	8	2	10	4
32	(3S)Galb1-4(Fuca1-3)GlcNAc-Sp0	7	11	16	8	10	17	6	32	17

302	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	142	12	8	6	3	4	14	10
235	Neu5Aca2-3Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	7	5	6	8	7	6	5	12	14
131	Fuca1-4(Galb1-3)GlcNAcb-Sp8	7	74	13	6	10	13	-2	30	18
598	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	7	19	17	1	7	12	4	15	15
263	Fuca1-2Galb1-4(6S)Glcb-Sp0	7	30	17	9	12	4	1	35	14
605	GlcNAcb1-6(Neu5Aca2-3Galb1-3)GalNAca-Sp14	7	12	6	2	22	9	1	6	18
403	GalNAca1-3GalNAcb1-3Gala1-4Galb1-4GlcNAcb-Sp0	7	20	11	6	8	13	-1	11	14
341	GlcNAca1-4Galb1-3GlcNAcb-Sp0	7	13	10	11	7	12	2	10	16
371	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6(GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	7	10	9	11	7	9	-1	16	7
180	GlcNAcb1-3GalNAca-Sp8	7	13	4	16	1	7	2	4	17
347	Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	7	10	11	9	5	12	2	10	15
424	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-3GalNAc-Sp14	7	232	3	4	23	3	1	21	27
66	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb-Sp8	7	17	8	3	10	5	5	24	9
421	Fuca1-2Galb1-4GlcNAcb1-2Mana1-6(Fuca1-2Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	7	14	11	**50**	7	11	0	23	11

272	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	7	13	10	5	23	5	2	17	7
409	GalNAcb1-3Gala1-6Galb1-4Glcb-Sp8	7	5	17	6	9	13	-3	14	9
73	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb-Sp8	7	17	10	3	9	7	-3	13	11
525	Galb1-3GlcNAcb1-2Mana-Sp0	7	255	30	6	17	14	1	40	26
212	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	11	14	9	7	11	1	13	11
374	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6(GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	7	9	6	6	12	3	1	13	5
318	Neu5Aca2-3Galb1-4GlcNAcb1-6(Neu5Aca2-3Galb1-3)GalNAca-Sp14	7	20	8	10	4	17	2	5	3
455	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6(GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	6	9	19	12	5	13	3	12	18
328	Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6(Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	6	104	6	2	3	4	3	16	11
401	Fuca1-2Galb1-4GlcNAcb1-3GalNAca-Sp14	6	11	4	7	15	12	-1	21	23
201	G-ol-Sp8	6	14	14	16	8	7	3	18	16
499	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-6GalNAca-Sp14	6	26	30	10	14	3	3	11	19
423	Fuca1-2Galb1-3GlcNAcb1-3GalNAc-Sp14	6	13	27	7	10	7	0	6	14
295	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	6	11	13	5	1	1	-2	0	6

96	GalNAcb1-3(Fuca1-2)Galb-Sp8	6	13	8	10	5	6	4	12	7
331	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	6	12	9	2	7	9	2	6	10
310	GlcNAcb1-3Man-Sp10	6	15	13	2	8	1	1	19	24
532	Galb1-4GlcNAcb1-2 Mana1-6(GlcNAcb1-4)(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAc-Sp21	6	944	25	-2	3	5	2	32	23
530	Neu5Aca2-3Galb1-3GlcNAcb1-4Galb1-4Glcb-Sp0	6	3	9	3	12	3	-2	2	3
567	GlcNAb1-3Galb1-3GalNAc-Sp14	6	11	11	1	10	6	0	16	5
239	Neu5Aca2-3Galb1-3(6S)GlcNAc-Sp8	6	24	29	7	8	14	-3	15	24
288	Neu5Gca-Sp8	6	20	13	7	15	7	-1	19	7
267	Neu5Aca2-6GalNAcb1-4GlcNAcb-Sp0	6	17	4	2	8	4	-1	13	4
286	Neu5Gca2-6GalNAca-Sp0	6	8	4	3	19	6	2	11	12
470	Glca1-6Glca1-6Glca1-6Glcb-Sp10	6	21	9	7	14	18	-1	18	9
240	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp8	6	19	11	10	10	9	2	16	8
271	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	6	10	9	2	13	8	-1	8	8
505	(3S)GalNAcb1-4(Fuca1-3)GlcNAcb-Sp8	6	25	15	7	6	4	-2	10	5
452	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-6(GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3)GalNAc-Sp14	6	25	13	7	6	2	-1	9	4
459	Galb1-4GlcNAcb1-6(Galb1-4GlcNAcb1-2)Mana1-6(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	6	2129	117	2	11	5	1	60	33

606	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-6(Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3)GalNAca-Sp14	6	10	20	10	20	**28**	3	16	11
188	GlcNAcb1-6(GlcNAcb1-4)GalNAca-Sp8	6	10	9	4	18	12	-2	4	7
513	(6S)GalNAcb1-4GlcNAc-Sp8	6	23	7	2	10	8	-1	17	13
555	Neu5Gca2-8Neu5Aca2-3Galb1-4GlcNAc-Sp0	6	24	13	5	5	5	4	15	5
608	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	6	24	24	14	23	10	5	25	25
432	GlcNAcb1-2Mana1-6(GlcNAcb1-4)(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	6	2	6	14	10	6	0	4	11
495	Fuca1-2Galb1-4GlcNAcb1-6GalNAca-Sp14	6	49	19	11	8	3	2	13	9
108	Gala1-3(Fuca1-2)Galb-Sp8	6	803	15	6	10	9	-1	74	5
91	GalNAca1-3(Fuca1-2)Galb-Sp18	6	10	11	22	9	7	1	9	14
139	Neu5Aca2-6(Galb1-3)GlcNAcb1-4Galb1-4Glcb-Sp10	5	23	16	-1	4	1	-3	8	14
390	Fuca1-2Galb1-3GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	5	12	14	1	2	4	2	17	8
192	GlcNAcb1-4GlcNAcb1-4GlcNAcb-Sp8	5	9	9	8	9	0	-2	12	9
105	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp8	5	121	9	2	9	14	-1	19	8
519	(6P)Galb1-4GlcNAcb-SP0	5	15	9	10	13	13	0	10	4
593	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	5	14	7	20	6	**26**	1	18	6

434	GlcNAcb1-6(GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp21	5	8	5	6	-2	6	3	7	12
213	Mana1-2Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana1-2Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	5	14	16	5	23	1	5	15	9
227	GalNAcb1-4(Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3)Galb1-4Glcb-Sp0	5	8	11	17	20	1	1	17	14
177	GlcNAcb1-6(GlcNAcb1-3)GalNAca-Sp8	5	8	15	3	10	4	1	11	10
454	GalNAcb1-4Galb1-4Glcb-Sp0	5	29	13	11	8	14	3	26	18
359	KDNa2-3Galb1-4Glc-Sp0	5	17	15	11	5	3	2	14	15
570	(3S)GlcAb1-3Galb1-4GlcNAcb1-3Galb1-4Glc-Sp0	5	10	7	13	10	-1	-2	8	34
442	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb-Sp8	5	21	12	3	7	10	0	12	5
184	GlcNAcb1-3Galb1-4GlcNAcb-Sp8	5	15	10	5	3	4	1	17	6
399	Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp12	5	56	10	4	2	1	2	9	16
40	(4S)Galb1-4GlcNAcb-Sp8	5	9	8	3	12	9	-1	17	30
512	(6S)(4S)GalNAcb1-4GlcNAc-Sp8	5	19	3	-1	7	18	-2	7	7
336	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	5	6	8	7	4	7	-2	13	9
241	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	5	16	24	20	0	5	0	21	25
377	Neu5Aca2-3Galb1-4GlcNAcb1-3GalNAc-Sp14	5	32	6	8	10	14	-1	15	5

599	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	5	9	22	2	23	21	-4	9	3
193	GlcNAcb1-6GalNAca-Sp8	5	13	10	12	11	11	5	10	32
217	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	5	18	25	0	12	3	3	13	2
238	Neu5Aca2-3GalNAcb1-4GlcNAcb-Sp0	5	8	10	12	9	6	2	18	14
550	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp25	5	30	15	10	8	10	0	24	4
231	Neu5Aca2-3(6S)Galb1-4(Fuca1-3)GlcNAcb-Sp8	5	15	7	12	8	-2	3	9	0
416	GalNAca1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	5	1	16	5	5	7	0	15	8
533	Galb1-4GlcNAcb1-2 Mana1-6(Galb1-4GlcNAcb1-4)(Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAc-Sp21	5	228	28	5	9	5	0	14	28
266	Neu5Aca2-6GalNAca-Sp8	5	11	6	4	5	11	0	10	15
531	GlcNAcb1-2 Mana1-6(GlcNAcb1-4)(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAc-Sp21	5	17	11	1	7	6	1	5	25
537	GalNAca1-3(Fuca1-2)Galb1-3GalNAcb1-3Gala1-4Galb1-4Glc-Sp21	5	9	5	11	4	0	-1	5	16
89	GlcNAcb1-3Galb1-3GalNAca-Sp8	5	9	13	5	14	11	1	15	8
77	Fuca1-2Galb1-4GlcNAcb-Sp8	5	27	6	4	5	9	1	12	10

259	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	5	6	14	4	8	14	3	14	9
276	Neu5Aca2-8Neu5Aca-Sp8	5	5	14	8	11	1	-1	19	5
51	Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	4	30	12	9	9	1	0	9	17
529	GalNAcb1-4GlcNAcb1-2Mana-Sp0	4	9	5	14	13	2	-1	12	8
65	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp9	4	5	19	7	10	7	-2	14	20
460	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6(GlcNAcb1-4)(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	4	4	8	8	3	6	2	7	13
103	Gala1-3(Fuca1-2)Galb1-3GlcNAcb-Sp8	4	1036	3	3	5	6	5	43	8
166	Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp8	4	168	80	12	13	9	3	28	27
536	GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3)Galb1-4GlcNAc-Sp0	4	10	0	6	8	7	1	11	20
544	Neu5Gca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Gca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp24	4	23	25	14	9	13	2	17	17
56	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	4	24	21	13	18		-5	18	21
485	Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp19	4	8	10	2	23	7	3	7	14
255	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	4	16	5	1	6	5	1	29	10
557	Neu5Gca2-8Neu5Gca2-6Galb1-4GlcNAc-Sp0	4	44	9	7	15	8	4	15	15

219	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	4	14	8	13	8	1	-1	7	12
574	Neu5Aca2-8Neu5Aca2-3Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glc-Sp21	4	17	8	15	11	23	-1	5	7
236	Neu5Aca2-6(Neu5Aca2-3)GalNAca-Sp8	4	8	4	1	10	16	0	9	5
186	GlcNAcb1-3Galb1-4Glcb-Sp0	4	9	12	0	7	17	-1	9	12
174	GlcNAca1-3Galb1-4GlcNAcb-Sp8	4	17	22	4	11	3	2	8	9
482	Neu5Aca2-6Galb1-4 GlcNAcb1-6(Neu5Aca2-6Galb1-4GlcNAcb1-3)GalNAca-Sp14	4	22	19	7	8	9	2	18	8
274	Neu5Aca2-6Galb1-4Glcb-Sp8	4	6	14	14	13	13	6	9	3
426	Gala1-3Galb1-3GlcNAcb1-3GalNAc-Sp14	4	**3921**	11	**71**	6	8	1	435	12
364	Gala1-3Galb1-4GlcNAcb1-2Mana1-6(Gala1-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	4	**7906**	29	5	8	9	1	**2103**	14
334	Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	4	**8463**	47	8	16	14	4	**2343**	12
75	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	4	15	9	12	10	8	-1	10	19
508	GlcNAcb1-6(GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(GlcNAcb1-4(GlcNAcb1-2)Mana1-3)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAc-Sp21	4	5	8	10	12	13	-5	12	4
556	Neu5Gca2-8Neu5Gca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAc-Sp0	4	13	7	11	7	3	2	10	12
469	Gala1-3(Fuca1-2)Galb1-3GalNAcb-Sp8	3	58	15	15	14	6	-1	22	27
183	GlcNAcb1-3Galb1-4GlcNAcb-Sp0	3	-1	19	3	14	4	2	0	6

358	KDNa2-6Galb1-4GlcNAc-Sp0	3	11	7	7	3	2	-1	16	14
53	GlcNAcb1-2Mana1-6(GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	3	17	8	7	5	4	1	4	5
90	GalNAca1-3(Fuca1-2)Galb-Sp8	3	8	20	7	13	15	2	7	8
408	Galb1-3GlcNAca1-6Galb1-4GlcNAcb-Sp0	3	56	19	6	6	5	5	10	5
516	(3S)GalNAcb1-4GlcNAc-Sp8	3	7	15	9	13	-1	1	13	6
596	GlcNAcb1-3Galb1-4GlcNAcb1-6(GlcNAcb1-3Galb1-4GlcNAcb1-3)GalNAca-Sp14	3	7	22	18	4	2	-2	10	17
297	4S(3S)Galb1-4GlcNAcb-Sp0	3	11	15	3	12	9	-1	12	9
563	GalNAcb1-3GlcNAcb-Sp0	3	7	18	5	18	9	-1	5	7
199	Glcb1-4Glcb-Sp8	3	20	26	8	23	10	0	22	32
312	GlcNAcb1-4GlcNAcb-Sp12	3	9	23	3	6	3	3	16	2
466	Neu5Aca2-6Galb1-4GlcNAcb1-6(Neu5Aca2-6Galb1-4GlcNAcb1-2)Mana1-6(GlcNAcb1-4)(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	3	17	6	6	15	9	1	17	13
326	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	3	7	20	6	6	11	1	7	6
558	Neu5Aca2-8Neu5Aca2-3Galb1-4GlcNAc-Sp0	3	13	6	4	4	19	-2	7	9
153	Galb1-4(Fuca1-3)GlcNAcb-Sp8	3	108	19	5	9	5	2	19	7
523	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana-Sp0	3	665	11	-2	9	11	-1	31	16
216	Mana1-6(Mana1-3)Mana1-6(Mana1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	3	22	9	12	5	9	0	15	5

319	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	3	10	8	9	1	6	-2	12	3
208	Mana1-2Mana1-2Mana1-3Mana-Sp9	3	14	18	3	4	6	1	20	16
202	GlcAa-Sp8	2	29	17	5	7	12	0	18	8
496	Gala1-3Galb1-4GlcNAcb1-6GalNAca-Sp14	2	**5164**	18	**38**	13	12	3	**954**	8
262	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	2	11	6	8	**30**	12	2	14	13
527	Neu5Aca2-3Galb1-3GlcNAcb1-2Mana-Sp0	2	9	6	5	6	15	2	5	7
481	Neu5Aca2-6Galb1-4GlcNAcb1-6GalNAca-Sp14	2	20	20	9	25	3	2	18	10
257	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb-Sp8	2	2	10	1	6	7	4	17	1
507	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-6GalNAca-Sp14	2	17	14	17	11	16	2	24	6
594	GlcNAcb1-3Galb1-4GlcNAcb1-3GalNAca-Sp14	2	5	16	5	4	9	2	6	8
351	Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	2	53	14	7	1	12	-2	22	11
281	Neu5Gca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	2	5	7	2	13	14	-2	14	13
348	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	1	17	13	11	6	3	5	3	5
118	Gala1-3Galb1-4Glc-Sp10	1	**3974**	18	**36**	10	**24**	0	**602**	17
397	Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-6(Neu5Aca2-3Galb1-3GlcNAcb1-2Mana1-3)Manb1-4GlcNAcb1-4GlcNAc-Sp19	1	18	4	7	7	10	3	11	8
179	GlcNAcb1-6(GlcNAcb1-3)Galb1-4GlcNAcb-Sp8	1	13	16	21	9	16	4	9	12

185	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	1	14	15	8	24	-2	-1	6	15
333	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	1	6	10	3	11	8	3	9	3
431	Fuca1-3GlcNAcb1-6(Galb1-4GlcNAcb1-3)Galb1-4Glc-Sp21	0	140	35	8	4	12	2	23	10
277	Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb-Sp0	0	4	23	4	4	10	-2	16	9
256	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp8	0	9	4	4	2	5	2	19	7
522	Gala1-3Galb1-4GlcNAcb1-2Mana-Sp0	-1	3731	19	**29**	7	10	2	**915**	10
87	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	-2	14	7	6	4	8	3	9	0
279	Neu5Acb2-6GalNAca-Sp8	-7	3	-3	4	7	4	3	12	-7
114	Gala1-3Galb1-4(Fuca1-3)GlcNAcb-Sp8	-8	2245	2	10	7	12	2	166	-7

Intended to be blank

Chapter 3 Implications of NAD^+ limitation and a *BAS1Δ*

3.1 Introduction

Proper maintenance of cellular concentrations of nicotinamide adenine dinucleotide (NAD^+) are critical for a multitude of important metabolic and biological processes such as (DNA repair, transcription, redox state regulation, Ca^{2+} stores, apoptosis, stress resistance and endocrine signaling) and when imbalances occur they often lead to various human diseases or even death as in the case with pellagra; a rare occurrence today associated with malnutrition.[52] Vitamin B3 (niacin/nicotinic acid (NA) and its derivatives including the reduced and phosphorylated forms (NAD^+, NADH, $NADP^+$, and NADPH) are classically known as coenzymes that participate and function as hydride acceptors and donors in a variety of cellular redox reactions.[52–54] In redox reactions, the NAD^+/NADH and $NADP^+$/NADPH pools are recycled and interconverted but overall concentrations are essentially stable. Additionally, NAD^+ can also act as a substrate that is consumed by NAD^+ cleaving enzymes such as ADP-ribose transferases, cADP-ribose synthases and Sirtuin-type (Type III) protein deacetylases.[53] These reactions act to decrease the NAD^+/NADH pool and require a constant resynthesis of NAD^+ either by salvage of NAD^+ catabolites or through de-novo synthesis. The Cormack lab has previously shown that sirtuins in the yeast pathogen *C. glabrata* consume and deplete NAD^+ reserves and

that reduced levels of NAD^+ can regulate virulence by rendering the NAD^+ -dependent Sir2p histone-deacetylase (HDAC) inactive which in turn results in transcriptional derepression of some Epa-adhesin (Epa6, Epa7) proteins critical for adherence to epithelial cells and for murine urinary tract infections (UTI).[16]

Some of the many interesting features of the yeast pathogen *C. glabrata* relate to its evolutionary history. Whole genome duplication events followed by extensive gene loss (such as the loss of genes necessary to catabolize galactose and allantoin) and possible adaptations during human host coevolution has rendered it auxotrophic for many metabolically important pathways known to exist in its evolutionary related yeasts – *S. cerevisiae* and in the most common human yeast pathogen *C. albicans*.[55] This creates a survival imperative to function as a commensal and/or as a pathogen for the acquisition of essential micronutrients (pyridoxine, thiamine, nicotinic acid).[56] While *C. glabrata* and *S. cerevisiae* are phylogenetically closely related, *S. cerevisiae* remains non-pathogenic and retains its capacity to synthesize NAD^+ *de novo*, by contrast, the pathogenic *C. glabrata* has eliminated the genes encoding the kynurenine pathway, rendering it unable to synthesize NAD^+ *de novo*, meaning therefore, that environmental limitation for NA in *C. glabrata* leads to cellular NAD^+ depletion. Growth limiting experiments for other micronutrients such as – pyridoxine and thiamine have shown

that only NA limiting results in the specific induction of *EPA6*.[16] Thus, the data are consistent with a model in which the NAD^+ dependent sirtuins lose functionality under NA limiting conditions.[57,58] In addition to the previously uncovered results on *EPA* gene expression mediating adhesion to epithelial cells, NAD^+ limitation also results in an increase in virulence in *C. glabrata*. In animal studies, following delivery of the inoculum into the mouse blood stream, fungal burden is subsequently assessed 7 days post infection. Infection with NAD^+ starved cells results in a fungal burden 10-100x greater than seen with cells that are harvested in log phase or cells grown until stationary phase conditions. Previous work to try and identify and delineate the key regulators and signaling pathways involved in the niacin limitation response led to experiments using in-house microarrays to obtain an transcript expression profile under such NA-limiting (virulent) conditions. The results indicate, as expected, that the majority of the genes induced are regulated by the sirtuins – Hst1 and Sir2.[59] Surprisingly, the effect of NAD^+ limitation on increased virulence is not mediated by these key regulators – i.e. the NAD^+-dependent regulators Hst1 or Sir2, since the same effect on increasing virulence by limiting NAD^+ was seen with *hst1*Δ *sir2*Δ mutant strains. There must therefore be at least one additional signaling pathway responsible for the effect on virulence. These findings are the impetus driving the following studies.

To analyze the transcriptional response to NAD^+ limitation and explore potential pathways important in regulating virulence, we first used microarrays to gather a genome-wide response to NAD^+ limitation. Within the several hundred genes induced by NA limitation, large sets of strongly up regulated genes are orthologues of *S. cerevisiae* genes regulated by the transcription factors Bas1 and Bas2. In *S. cerevisiae*, *BAS1* and *BAS2* regulate several interconnected metabolic pathways, including histidine, adenine biosynthesis, THF (tetrahydrofolate) metabolism, and phosphate utilization.[60–62] We analyzed the transcriptional response of the Bas1 and Bas2 regulons in *C. glabrata* in response to NAD^+ limitation. Essentially, almost all the genes associated with purine biosynthesis and the interconnected pathways are transcriptionally activated by NAD^+ limitation. As a follow up to these initial findings we extended the microarray analysis in the context of mutants lacking the *bas1* transcription factor. Our results indicate compromised expression of key genes representative of the above mentioned pathways. We also show that the mutant strain *bas1*Δ is highly avirulent, consistent with the idea of *BAS* genes being potential regulators of the virulence in *C. glabrata*.

The *BAS1* and *BAS2* (basal) genes are recognized as global transcriptional regulatory proteins.[63] More specifically they were initially identified as basal regulators of *HIS4* epression.[64] In *S. cerevisiae* there

are two control systems for the *HIS4* gene: general and basal. General control is a global response to amino acid starvation that is mediated by the transcription factor Gcn4 (general control non-derepressible).[65,66] Basal control or in the absence of amino acid starvation high basal level Gcn4-independent transcription of *HIS4* is maintained by the Bas1 and Bas2 proteins.[67] The extensive homology between S. *cerevisiae* and *C. glabrata* from phylogenetic studies reveals a high degree of syntenic organization which underlies the conservation of a large number of regulatory pathways.[55] Efficient utilization of cellular nutrients has allowed yeast cells to create very sensitive networks that operate at various interconnected levels. A sophisticated level of complexity has evolved in yeast for the efficient coordination and coregulation of diverse metabolic pathways like purine and histidine biosynthesis.[60] Here we find multiple points of convergence linking these pathways in a compelling way that support results from our qRT-PCR and (–NA) microarray analysis. A review of *S. cerevisiae* literature revealed that the transcription factors Bas1 and Gcn4 share common substrates and intermediates, such as: the genes (*ADE3, ADE4, HIS4, HIS7, SHM2*)[60,68,69] and the metabolites 10-formyl THF and AICAR – 5'phosphoribosyl-4-carboxamide-5-aminoimidazole.[60,70] Moreover, they both act through the same promoter motif 5' TGAC/GTC 3' although the flanking ends may differ somewhat.[68,71,69] Double deletion mutant strains of *gcn4*Δ and the *BAS* genes show a loss of expression activity under NA limiting

conditions that is in keeping with a synergistic or additive effect and consistent with the *bas1*Δ and (*bas1 bas2)*Δ mutants.

Our previous studies (Cormack-lab) dealt with the exploitation of host nutrient availability leading to NAD^+ depletion as a consequence of consumption specifically by sirtuin-HDACs resulting in the de-repression of sub-telomeric associated *EPA* adhesin genes. The following work addresses yet another aspect of NAD^+ limitation strictly relevant to its role in redox reactions and *C. glabrata* hypervirulence. Once exposed to limiting concentrations of NA, *C. glabrata* transitions from a benign commensal organism to an opportunistic yeast pathogen. The impact of the micronutrient starvation is so profound that normally repressed genes when in the presence of Adenine are bypassed, resulting in increased flux through the purine biosynthetic pathway including the interconnected pathways of histidine and tetrahydrofolate biosynthesis, that ends with the generation of AMP and GMP nucleotides. An analysis of the pathway affirmed the role of NAD^+ as a cofactor in IMP conversion. Remarkably, the hypervirulent phenotype observed under limiting NA conditions during murine disseminated infection experiments is replicated through the use of Mycophenolic acid (MPA), an inhibitor of IMP dehydrogenase. At one level we show that the transcriptional regulators Bas1 and Gcn4 are indirectly associated with *C. glabrata* virulence. Furthermore, and perhaps more meaningful are the results

which suggest that the driving force behind *C. glabrata* infection may be due to a direct impact on GMP production by an inability to convert IMP to XMP due to compromised levels of the enzymatic cofactor NAD^+ that renders inactive the oxidoreductase – IMP Dehydrogenase which ultimately leads to an imbalance of nucleotide pools and creates a constitutively active loop on central metabolism pathways in an effort to return to nucleotide pool homeostasis.

3.2 Results

3.2.1 NAD^+ limitation regulates transcription independently of HST1 and SIR2

The working model for gene activation by nicotinic acid (NA) via the sirtuins goes as follows: during periods of NA limiting environmental supply leads to a drop in the cellular NAD^+ concentration thereby reducing the function of the NAD^+ dependent histone deacetylases (HDACs) – the Sirtuin family of proteins. This eventually results in transcriptional derepression at normally silenced sub-telomeric regions particularly Hst1 and Sir2-repressed loci. Previous transcriptional profiling work in the Cormack lab determined that the primary sirtuin that mediates the transcriptional response of *C. glabrata* to NAD^+ limitation is Hst1, a Sir2 yeast homolog. Additionally, Hst1 was shown to regulate the *TNR1*, *TNR2* and *TNA1* genes encoding various NAD^+

precursors transporters.[59] Thus surprisingly, in a (*hst1,sir2*)Δ mutant the hyper-colonization phenotype remains intact as evidenced from results in a disseminated infection model (data not shown). To extend these earlier observations we carried out a set of microarray analysis of the (*hst1,sir2*)Δ strain grown under limiting NAD^+ versus replete NAD^+ conditions the results of which ultimately show the signature set of low-NAD^+ responding genes unperturbed in the mutant (Table 3.1). The summarized values (log_2 transformed) represent the average fold changes for two dye-swapped biological experiments. The blue highlighted numbers show the NAD^+ dependent transporters as well as Epa6 are Hst1, Sir2 regulated. The remaining sets of representative genes in the table belong to the three biosynthetic pathways (purine, histidine, one-carbon metabolism), which have been consistently validated by quantitative RT-PCR (Figure 3.1C, Figure 3.2B). Microarray gene induction values for *ADE1, HIS4, MTD1* and *SHM2* between wild type and (*hst1,sir2*)Δ are almost identical and again show similar expression levels in qPCR analysis discussed below. Other notable genes not analyzed by qPCR: *ADE4, ADE8, HIS7, HIS1* also remain responsive to limiting NA, apparently independent of Hst1 and Sir2 (Table 3.1).

One concern with the analysis above is that the NA-limited cells have stopped growing, while the NA-replete cells are in mid log phase. We wanted to confirm that the induction of the genes from the purine,

histidine and THF pathways was due to NA limitation rather than as a result of growth inhibition or a difference in cell densities (OD). Since we know that limitation for glucose has no effect on virulence in *C. glabrata* (i.e. there is no increase in kidney colonization, as seen with NA limitation), we chose to grow *C. glabrata* under limiting glucose conditions as a control. *C. glabrata* cells in NA replete SC media but with limiting glucose (0.04% dextrose) stop growing at the same OD_{600} as cells grown in limiting NA conditions (Figure 3.1A,B). Yeast cells were harvested along stationary phase growth for both NA starved and NA replete conditions (OD_{600} 1.5). RNA for qRT-PCR analysis was prepared as detailed in Materials and Methods. RNAs were normalized to HHT2 and expression induction ratios were determined. The (*hst1 sir2*)Δ strain shows that *TNR1* transcriptional induction was reduced 15-fold and transcriptional induction of *TNA1* was similarly reduced over 10-fold while *EPA6* was knocked down 3-fold, when compared to wild type expression levels (Figure 3.1C). This is expected as their regulation by Hst1 and Sir2 has been thoroughly established. Of significant consequence is the fact that our signature set of genes: *MTD1*, *SHM2*, *HIS4* and *ADE1* remained responsive with almost identical expression ratios compared to wild-type cells grown in limiting NA SC media (Figure 3.1C). These qPCR results along with mouse infection results where (*hst1, sir2*)Δ remains hypervirulent (data unpublished) give strong support to conclude that the low NA virulent signal is mediated

independently of the well known NAD^+ dependent regulators Hst1 or Sir2.[57,58]

3.2.2 The BAS transcription factors (Bas1, Bas2) activate transcription of target genes in response to limiting NAD^+

The limiting NA effect is characterized by growth inhibition evidenced by 24-hour growth monitoring (Figure 3.2A) of NA (3.25 µM) replete and NA (0.01625 µM) starved *C. glabrata* cells. A subset of genes that are highly up regulated in response to low NAD^+ levels strikingly belong to three distinct biosynthetic pathways. Interestingly, these three pathways (purine, histidine, single-carbon metabolism) display cross-pathway interdependent regulation.[60,61,70] Previous work (Cormack-lab) has found that almost every member of each of these pathways appears to be strongly induced. These pathways have been extensively studied and are known to be under control of the Bas1 and Bas2/Pho2 (basal) global transcription factors. Utilizing a series of deletion strains: *bas1*Δ, *bas2*Δ or the (*bas1 bas2*)Δ mutants we show that the absence of these transcription factors appears to attenuate the hyper-colonization phenotype as demonstrated by our murine disseminated infection experiments (Figure 3.2 C-D). Yeast cells for mouse infection experiments were prepared as detailed in Materials and Methods section and colony forming units (CFUs) of *C. glabrata* cells recovered from mouse target organs (kidneys) on Pen/Strep plates were assessed 2-days following

30°C incubation. As expected, BG2-wt cell counts for –NA displayed the hyper-colonization phenotype resulting in CFUs of over an order of magnitude greater than the non-starved (+NA) cells. Surprisingly, the *bas1*Δ mutant grown in limiting NA conditions did not show hyper-colonization. In fact the CFUs did not simply show avirulence but the counts were well below the +NA replete grown cells, almost a whole order of magnitude and curiously the *bas1*Δ +NA CFUs were significantly lower than the wild-type +NA CFUs. The *bas2*Δ mutant showed an intermediate effect not as clear or profound as *bas1*Δ while the (*bas1 bas2*)Δ recapitulates the *bas1*Δ effect (Figure 3.2C-D). The results suggest that Bas1 plays an important role in *C. glabrata* infection inhibition. Nonetheless, a concern arises when considering the +NA control group results for all the Bas mutants in both experiments. Even when the *bas1* mutants were not starved, their colonization of organs was lower than that seen for wild type cells. This raises the possibility that *BAS1* may be required for growth in vivo, as opposed to response to the NA depletion *per se*. This possibility prompted us to create a *bas1* conditional knockout – a description of those experiments and results are discussed further below.

Interestingly, the pre-conditioning of *C. glabrata* cells in low NA media allows the pathogen to overcome the Adenine repression, which results in increased activation thru the purine biosynthetic pathways as

well as the establishment of a virulent program. We explored the extent of the impact of the *BAS* genes in pathway regulation and *C. glabrata* hyper-virulence in the context of a limiting NA environment. Gene expression analysis utilizing representative genes of the three biosynthetic pathways was conducted by real-time quantitative reverse transcription PCR (qRT-PCR). Yeast cells were harvested simultaneously at matching times (5hrs) and OD_{600} (~0.5) in early log phase for qPCR analysis. Our mouse infection results are further supported by the expression inhibition of our signature low NAD^+ responsive genes – as shown in the qPCR results (Figure 3.2B). The extent of NAD^+ starvation is indirectly shown by the strong induction of the nicotinic acid transporters (*TNR1, TNA1*) while the *bas1*Δ and *bas2*Δ mutant strains ratios show an almost 10 fold reduction in the level of expression for most of our three pathway representative genes tested: *MTD1*, *SHM2*, *ADE1*. Notably, the histidine gene (*HIS4*) is relatively unaffected and appears to behave independent of the Bas factors. Although our results strongly suggest a role for Bas1 inhibition, we can see that the *bas1*Δ alone fails to activate genes associated with central metabolism; however, the result is confounded due to the potential slow growth effect. Others have reported a slow growth phenotype for the *bas1*Δ although this has been mainly attributed to the absence of adenine.[60,68] We have examined *bas1*Δ growth conditions for multiple *bas1*Δ isolates (#2521, #2522,

#3354) both in SDC and YPD media and have found no slow growth effects for the *bas1*Δ mutant (Figure 3.3 A-B).

The limiting NAD^+ response is not unique to the clinical isolate BG2 and has also been catalogued for the CBS138 *C. glabrata* strain (Figure 3.3C). qPCR analyses also show comparable levels of gene expression as seen in the BG2 strain. To help establish a sense of our signature genes transcript expression levels we carried out a time course analysis of –NA induction across a 24-hour period by qRT-PCR. Expression levels were normalized to log phase (OD_{600} =0.5) +NA cells. On average, peak expression level ratios appear to max-out around 8-10 hours (Figure 3.3D).

Moreover, the microarray analysis of the *bas1*Δ strain provides compelling evidence that additional factors could be involved in the regulation of the interlinked biosynthetic pathways of purine, histidine and one-carbon metabolism. *C. glabrata* cells were grown overnight in SDC media and excess NA (10 μM) to an optical density (OD_{600}) of ~ 0.5. The following day the culture was re-inoculated into SDC media in the presence of 3.25 μM NA and SDC without NA supplementation at a starting OD_{600} of 0.05. Yeast cells were recovered from log phase OD_{600} ~ 0.5 and processed for RNA extraction and preparation to be used in microarray analysis. Table 3.2 summarizes a list of genes representing

the three up-regulated biosynthetic pathways discussed earlier. Shown is a summary of log_2-transformed values representing the average fold changes for two microarray experiments. During +NA replete conditions most genes were not affected by the *bas1*Δ; as expected. One gene did stand out with high basal level expression, namely *MTD1*. When analyzing the mutant effect under induced (-NA) conditions most of the genes failed to be expressed, again as expected, except for several of the *HIS* regulon members, indicating possible Bas1 independent activation. The BG2 wild type induced up-regulation of our listed genes was similar, reproducible and consistent with previous data (Cormack-lab). The final column of the table illustrates the induction capacity (-NA vs. +NA) of the *bas1*Δ mutant. Surprisingly, about half of the genes appear to be compromised for expression while the other half show variable up-regulated expression levels. Unexpectedly, the *bas1*Δ strain indicates residual activity, from *MTD1* and glaringly, the entire Histidine regulon (Table 3.2, blue highlighted values). Here, we show that the persistent activation of these genes might be ascribed to the presence some alternate factor (Bas1-independent). The induced genes from the *bas1*Δ array showed high-level expression that is comparable to results seen in BG2 – wild type. Many of the genes showed compromised activation levels while the activation of the entire *HIS* regulon and *MTD1* prompted closer investigation eventually leading to a cursory review of the literature that confirmed the participation of Gcn4 in all three

biosynthetic pathways (histidine, purine, 10-formyl THF metabolism).[62,64,69] Gcn4 is well know to be responsible for *HIS4* activation. These results are in keeping with published literature and conclude that Bas1 clearly plays a key role in the regulation of these pathways in the context of low [NAD^+] as well as in hyper-virulence to some extent. A Gcn4 role if any, in *C. glabrata* colonization remains to be established.

3.2.3 Gcn4 works additively with Bas1 to activate transcription in response to limiting NAD^+

Interestingly, Gcn4 and Bas1 share and bind independently to the same DNA responsive motif (TGACTC).[68,69] Gcn4 is a very well known global transcription factor that becomes translationally responsive under various stressful conditions such as amino acid deprivation.[65] Apart from its involvement in histidine biosynthesis, previous studies have shown that Gcn4 is also necessary for optimal expression of several purine genes (e.g. *ADE3, ADE4*). Our microarray data indicated the presence of additional regulatory factors acting in response to the limiting NA condition. Consequently, we created a set of *gcn4*Δ strains along with *bas1*Δ to explore the role of this factor during low NA exposure in *C. glabrata*. Preliminary work on the growth phenotype of the various deletion strains indicated the need for supplementation by Tryptophan in SC media (Figure 3.4). *C. glabrata* yeast cells were grown overnight in SC

+30 μM excess NA, then inoculated into SD CAA media. Growth curves of strains grown in casamino acid (CAA) media showed that a slow growth effect in Gcn4 mutants could be overcome by Tryptophan addition. Similarly, an extended lag phase for Bas deletions was eliminated by the addition of Adenine in SD CAA media.

Our data from qRT-PCRs of the *gcn4*Δ double knockout mutants uncovered an increased loss of activation for all of our three pathways representative genes in conjunction with the *bas1*Δ and *bas2*Δ deletions but no effect was observed for the *gcn4*Δ mutant alone. Yeast cells were harvested after 12hrs of NA starvation in SC +Ade, +Trp media versus log phase cells harvested at 6hrs in replete (+3.25 μM NA) SC media +Ade, +Trp. RNA was prepared for qPCR analysis as detailed in Materials and Methods. The RNAs from three separate biological experiments were normalized to *HHT2* and the averaged values were used to determine the induced gene expression ratios. For all genes tested: *MTD1*, *SHM2*, *HIS4*, *ADE1* there is a dramatic reduction in expression levels compared to BG2-wild type and a greater than 10-fold reduction of expression compared to the single *bas1*Δ gene deletion mutant. This additive effect raises the possibility that Bas1 and Gcn4 may function cooperatively under the stressful condition of NAD^+ starvation (Figure 3.5 A). Here, we show that Gcn4 alone does not have an apparent effect on the expression of our limited-NA responsive genes, the effect is present only with

double-deletions of the *BAS* genes that we see complete reduction of gene activity ranging from 50-100 fold loss of expression. Whether Bas1 or Bas2, and Gcn4 are able to form heterodimeric complexes has not been observed experimentally nor did we attempt to address this potential scenario. To provide further evidence of the (*gcn4, bas1/2*)Δ additive repressive effects we also performed qPCR analysis from OD matched yeast cells grown in SC +(3.25μM NA) and limiting glucose (0.04% dextrose) and recovered strains along early stationary phase grown in the absence of NA (Figure 3.5 B). This approach adds further evidence that the effect is strictly due to NAD^+ status and not to a growth phase difference, i.e. cell density (OD) changes. Similar results were observed.

Considering the suppressed transcripts results for the three biosynthetic pathways we looked to address the possibility that a Gcn4 mutant might attenuate the hyper-virulent program of *C. glabrata* much like has been observed by Bas1. Consequently, we performed murine disseminated infection experiments to assess the role of Gcn4. Yeast cells were prepared as described in Material and Methods and colony forming units (CFUs) were recorded from YPD (Pen/Strep) plates. After two experiments attempted, our results concluded that *gcn4*Δ does not affect *C. glabrata* hyper-colonization (Figure 3.5 C-D), although we have shown a convincing effect on gene expression for our purine pathways in

response to limiting NA conditions. This effect is strictly due to synergism with the *BAS* genes.

3.2.4 Bas1 conditional knockout mutant: (Bas1 – AID – flag3)

To unravel the possible confounding results of Bas1 virulence attenuation due to the slow growth effect of the *bas1*Δ mutant in a non-ambiguous manner we implemented a conditional knockout system for *BAS1*. Our strategy makes a single modest assumption. Just as *C. glabrata* cells are pre-conditioned to limiting [NAD^+] and yet display a hyper-colonizing phenotype days after the initial infection, then a temporary "pre-conditioning" of a *bas1* deletion strain should also render an attenuated avirulent phenotype. To clarify this experiment, *C. glabrata* cells to be used in the infection inoculum are grown in the presence of IAA (auxin) to swiftly degrade our target Bas1p. Once yeast cells are harvested they are washed 2xPBS before mouse injection. At this time any remaining Auxin ought to be beyond negligible allowing for efficient recovery of the Bas1p knockout consequently no slow growth effect should be evident. For this experiment we used a newly established Auxin-based degron system for the rapid depletion of proteins.[72] The advantage of the system rests on the uniqueness of the auxin compound as well as the Transport Inhibitor Response-1 (TIR-1) component that are found only in plants and that bear no known homologs in the yeast. The strategy of the system is supported by the fact that the SCF (skip, cullen,

F-factor) complex is evolutionary conserved and maintained from bacteria to humans. Briefly, once a protein of interest has been modified with the AID (auxin inducible degron) tag, auxin (IAA-indole-3-acetic acid) is added to your culture at the desired time where it promotes the interaction between TIR1 and the target protein as depicted in the schematic (Figure 3.6). Then SCF-TIR1 becomes active as an E3 ubiquitin ligase that eventually results in the poly-ubiquitination of the AID-degron leading to targeted clearance by the ubiquitin-proteasome pathway.

Prior to testing the Bas1 conditional knockout in mouse experiments, we performed a series of *in-vitro* tests of the (*Bas1 – AID – flag*$_3$) construct to ensure appropriate function and predictability of the AID system. It is worth mentioning some unsuccessful work regarding the initial design of this construct that placed a 3x Flag tag immediately after the *BAS1* ORF followed by the AID degron. This in fact results in an unstable, inoperative Bas1 protein. A redesign of the system places the 3xMiniDegron (219 amino acid - AID) in frame to BAS1 at the C-terminal genomic locus linked to a 3x – Flag tag for detection and a final Stop codon in the parental strain BG14 ($URA^{=}$). An unexposed AID tag is essential for proper system implementation. The oligos used in this cloning can be found in Table 3.3 and a description in the Material and Methods section. The fully functional AID system also contains the TIR1

component protein expressed in the pGrb2.3 URA+ plasmid. The Bas1-AID-flag$_3$ (#BG3857) strain was tested for the low growth phenotype observed when wild type strains are cultured in SC media –NA and for functionality regarding gene induction of our set of representative –NA responsive genes. The Bas1-AID-flag$_3$ strain showed a similar growth arrest response as the BG2-wild type when monitored in SC media without NA supplementation (Figure 3.7 A). Bas1-AID-flag$_3$ (BG3857), appears to function normally as similar activation levels for our signature set of –NA responsive genes in BG2-wild type -NA, were duplicated with at least 10-fold induction levels obtained in qRT-PCR results (Figure 3.7 B). Quantitative RT-PCR was also implemented to determine the transcriptional response or the loss of activation for a chosen number of genes when the conditional mutant system was fully tested with the addition of auxin. Yeast cells were grown overnight to log phase and then re-inoculated to an OD density of 0.025 in the presence or absence of NA. Auxin (IAA – 500 μM) was also added at this time to the –NA cells until they were harvested at 12 hours of NA starvation versus log phase (OD= 0.5) +NA cells. The qRT-PCR analysis reveals that there is a bas1-knockout effect for the genes tested: SHM2 and ADE1 (Figure 3.7 C-D). Although not very impressive there is a clear difference between the auxin+ versus the auxin- transcript level samples. We included some additional controls to guard against any possible influence that SDC – URA media might have on expression due to the parental strain BG 430.

Our results indicate that the Bas1-AID-Flag$_3$ +TIR1 +Auxin experiment mirrors the *bas1*Δ effect displayed by the loss of gene activation.

Additional tests were conducted for the *BAS1* conditional knockout. A growth curve test was made to determine if we could observe the same behavior of a *bas1*Δ extended lag phase when cells are grown in –Adenine conditions. Results of the growth curve (Figure 3.8A) show the knockout strain (BG3859 +IAA) directly overlaid onto a *bas1*Δ when Auxin (500 μM) is added to SDC media. More convincingly is the Western analysis of Bas1-AID-flag$_3$. Here we show that the AID system is extremely efficient and sensitive as a time course degradation analysis was carried out whereby samples were recovered at various times during auxin treatment: 20min, 40min, 1hr, 2hr and 4hours (Figure 3.8 B). Cell lysates were prepared for all three samples (BG3857, BG3859 +IAA, BG3859 –IAA) and 40 μg were run on 12% gradient PAGE denaturing gels for one hour. Subsequently, the gel was transferred onto a PVDF membrane and the Western blot was probed with an anti-Flag ab (1:2 K dilution, 2° 1:10 K dilution). The Bas1p is completely degraded by our earliest time point 20 minutes and is maintained at least through 4 hours. One final test checked for the stability of Bas1- AID – flag construct by growing the strain in the absence of Adenine. Three different Bas-AID-flag conditional strain isolates were analyzed and only the *bas1*Δ displayed an extended lag phase phenotype when grown in SC

–Ade (Figure 3.8C) suggestive of tight control over Bas1 degradation. We concluded that the AID system was stable and operational for Bas1p. Surprisingly, our murine disseminated infection experiment for the conditional knockout of Bas1p yielded a perfectly confounding result. The BG2 wild type behaved accordingly including a control +NA grown to low OD_{600} in limiting glucose (0.01% dextrose). The unexpected CFU counts for the Bas1-AID TIR1 strain approached –NA wild type levels (Figure 3.8 D). Moreover, the negative control with an empty pGrb2.3 plasmid and no TIR1 component shows very similar CFUs result that resembles the complete Bas1-knockout system. The results from the data imply that the effect at the micronutrient level is not mediated through Bas1 during the starvation phase. Assuming that the system did work then Bas1p was quickly restored and the cells were able to recover fully. Alternatively, considering that Bas1p is constitutively bound to DNA and is not up regulated during limiting NA exposure it is possible that the protein could be masked in a DNA/promoter complex that allows for some brief residual functionality; enough to activate our signature set of genes. Adding auxin earlier in the growth phase is not practical since that would likely introduce the slow growth effect we are directly trying to bypass.

3.2.5 IMPDH inhibition mimics the effect of limiting NAD^+ in C. glabrata

Noticing how responsive the central metabolism pathways are to the starvation condition of low [NAD^+] we investigated the possibility that there might be a direct link with NAD^+ along one of the three pathways that are strongly up regulated. Inspection of these pathways reveals in fact, NAD^+ does impinge at the branchpoint conversion of inosinate (IMP); the GMP arm of purine biosynthesis. NAD^+ functions as a coenzyme in the conversion of IMP to xanthylate (XMP) that eventually leads to GMP generation. NAD^+ is the cofactor that impacts the enzyme IMP-Dehydrogenase (IMPD). This is critically important since excess AMP production has no way of being converted back into GMP. Interestingly, Mycophenolic acid (MPA) is a drug that specifically inhibits GTP synthesis and is widely used to achieve immunosuppression as a strategy to prevent allograft rejection.[73] Our working premise is that the absence of NAD^+ leads to a block of GMP/GTP synthesis leading to the activation of the pathways as a way to restore and increase the end products yielding GMP/GTP. A similar scenario has been shown in *S. cerevisiae* with mutants along the second half of purine biosynthesis where deletions that block anywhere downstream of the point of synthesis for the metabolites SAICAR or AICAR leads to constitutive activation of the pathway while mutants blocking the first part of the

pathway leads to adenine auxotrophy.[74,75] Presumably, inhibition of IMPD should result in a feed forward loop for the purine biosynthetic pathway.

Analyzing the transcriptional readout of MPA blocked cells through qRT-PCR was the first test of the hypothesis. Some preliminary optimization tests were carried out to help determine an appropriate level of [MPA] for our analyses. A working concentration of 25 μg/mL [MPA] was added to BG2 log phase cells (OD_{600}= 0.5) resulting in a compromised growth phenotype similar to what is observed with –NA [≤0.025 μM], (Figure 3.9 A), although the cells eventually do mostly recover. We then determined the extent of gene activation by harvesting cells at two different time points (4hrs and 7hrs) post MPA introduction (Figure 3.9 B). We examined our usual cohort of genes and found an average 10-fold expression level difference that is similar to induction levels obtained during limiting NA conditions. This data suggests that MPA inhibition of IMPD mimics NAD^+ starvation and leads to an increase in purine pathway activity. Conversely, if a block in GMP/GTP generation results in our signature genes being up regulated then the activation should be relieved by an excess of guanine supplementation in the growth media. Physiological levels of guanine are typically 0.15 mM G; we tested two different concentrations of guanine (G): 0.3 mM G and 0.9 mM G. As expected the compromised growth phenotype was eliminated

at both concentrations even when using MPA at 50 µg/mL (Figure 3.9 C). A similar analysis was conducted for our pathways representative genes. Expression was measured by qPCR and the results were in keeping with the growth curve observation. Gene activation was induced by IMPD inhibition and substantially reduced by the presence of guanine at both 0.3 mM G and 0.9 mM G concentrations almost in a linear fashion – expression levels were down to approximate basal level transcription (Figure 3.9 D). All the genes tested: *MTD1, SHM2, ADE1, ADE3, ADE4* and *HIS4* were demonstrably suppressed in their activation. Curiously, the *TNR1* transporter control was also down although this is truly inconsequential considering that under limiting NA, its typical gene activation ranges between 100-1000 fold when measured by qPCR. These results indicate that the MPA block can be overcome by supplementing with guanine. In light of these results it is tempting to consider that the limiting NAD^+ response can potentially be suppressed by excess addition of guanine. We made several attempts to address this possibility but were unsuccessful. It is also worth noting that our experiments to introduce guanine into cells were not trivial. Guanine presents a difficulty of solubilization. It requires manipulations that range from high pH (1N NaOH) down to acidic levels for SC media. It is possible that these earlier attempts need to be revisited with an improved guanine availability protocol. Alternative plausible explanations for the failure of excess guanine to cover the limiting NA phenotype may simply lie with

the extent of action of the micronutrient (vitamin B3). NAD^+ plays a major role in a number of different critical pathways that possibly have not yet been uncovered in the pathogenesis of *C. glabrata*.

The second test of our hypothesis, described earlier, is to assess the ability of MPA treated *C. glabrata* cells to adopt a hyper-virulent program leading to hyper-colonization in a manner similar to what is observed during limiting NA preconditioning exposure. To that end, we carried out a murine disseminated infection experiment to explore the possibility of IMPD inhibition mimicking the low [NAD^+] response (Figure 3.10). Remarkably, our results confirmed the hyper-colonization phenotype by MPA inhibition (50 μg/mL). Taken together, our latest data suggests that the driving factor behind NAD^+ starvation is the creation of a nucleotide imbalance between AMP and GMP ultimately resulting in the pathogenesis of *C. glabrata*.

3.3 Discussion

C. glabrata presents a growing public health issue within infectious diseases. It has been generally recognized as a human commensal but more recently has emerged as an opportunistic pathogen.[2] *C. glabrata* shares a common ancestor with *S. cerevisiae*. Between the two, there exists a small number of genetic differences yet through a series of acquired attributes, (e.g. variable adherence capacity, drug resistance,

starvation resistance, high stress tolerance) have enabled this species to survive and thrive in a hostile environment.[56,76] The decision whether to undertake a commensal or pathogenic lifestyle is not defined by its auxotrophies, yet it has evolved to make use of one (nicotinic acid - NA) as part of a hyper-virulent program. Results from our findings indicate that the pathogenic state of *C. glabrata* is directly tied to central metabolism and that disruption along the 2nd half of the purine *de novo* biosynthesis pathway at points downstream of where the metabolites AICAR and SAICAR are generated can lead to a hyper-virulent phenotype. We now have two approaches that elicit the same observed phenotype; exposure to an NA limiting environment and inhibition of nucleotide synthesis (GMP). This is currently an overstatement since it remains to be seen whether inhibition of every point along the pathway will yield the same results. Nonetheless, the data provides encouraging results that warrant the pursuit of this line of questioning. Can the end products of purine biosynthesis AMP and GMP overcome the NAD^+ limiting phenotype? Will constitutively active flux through the pathway such as the use of a bas1-bas2 fusion heterodimer complex mimic the – NA hyper-colonization result? Can the overexpression of any single member of the Adenine regulon result in *C. glabrata* pathogenesis or is the hypervirulent program much more complex beyond the purine biosynthetic pathway? Does the imbalance of nucleotide pools somehow link up with a mechanism that indirectly allows for the derepression of

EPA adhesins? Our results here provide a starting point for addressing the downstream events following NAD^+ limiting exposure.

We have utilized a set of deletion strains (*bas1*Δ, *bas2*Δ, (*bas1 bas2*)Δ, *gcn4*Δ, (*gcn4 bas1*)Δ, (*gcn4 bas2*)Δ) to better characterize and delineate key factors or pathways responsive to the NA limiting virulent signal. We began our analysis with an effort to try and extend the initial findings of a (*hst1 sir2*)Δ mutant that does not play a role in the NA limiting phenotype. These are very well known HDACs that are functionally dependent on NAD^+. Unpublished murine disseminated infection experiments showed that the hyper-colonization result was persistent in the sirtuin deletion strains. Therefore, we conducted a microarray analysis of the (*hst1 sir2*)Δ to potentially uncover genes responsive to the NA limiting signal. Again the data revealed the same sets of genes being up regulated, thereby highlighting and emphasizing the importance of these pathways in *C. glabrata* pathogenesis, namely: histidine biosynthesis, purine biosynthesis and 10-Formyl tetrahydrofolate metabolism. There still remains a possibility for alternate drivers within this set of up regulated genes that remain unidentified. For example, there may be other 'NAD^+ consuming' enzymes like polyADP ribosyl-polymerases or ADP ribosyl-transferases involved in cellular signaling processes vital for survival yet to be uncovered.

However we continued to focus on the regulators of these pathways, in particular the *BAS* genes.

The *BAS* genes are recognized as global transcriptional regulators. Bas1 is constitutively bound to promoter specific DNA sequences (TGACTC) in an inactive state until recruitment of Bas2. Conversely, Bas2 is not bound to any Bas2 specific DNA elements until it heterodimerizes and is allowed to transcriptionally activate genes in concert with Bas1. There are a number of studies describing the activation and regulatory mechanism of Adenine dependent genes by Bas transcription factors. Typically adenine exerts an inhibitory influence on the Bas1 – Bas2 interaction either directly or with components of the transcriptional machinery that antagonizes the Bas proteins' activation function. It is only when Adenine is removed that we see increased accumulation of the purine pathway metabolites SAICAR and AICAR that release of the repressive state is achieved and the Adenine regulon is allowed to be expressed. Our results show that despite the presence of Adenine, exposure to very limiting concentrations of NA, will allow this repressive state to be overcome quite readily. Interestingly, various studies have shown that Bas1 – Bas2 complex formation is not due to expression level changes of either protein or by nuclear exclusion nor a repressive factor but rather a masking effect of activation domains for both proteins that somehow involves adenine. All of our SC media

experiments were conducted with excess adenine (150 μM) indicating that this inhibitory effect is overcome by an alternate positive unmasking effect on the Bas1 and Bas2 proteins. This unmasking event occurs presumably by a generally accepted mechanism by which gradual accumulation of the metabolites SAICAR and AICAR provide the activation signal for Bas 1/2 transcription factor activation, although a physical interaction is yet to be demonstrated.

A recent series of transcriptome analysis studies of conditions leading to GDP or GTP depletion revealed that the great majority of the genes (71) affected at least 2-fold were under the control of the Gcn4 transcription factor.[73] Hence, it was somewhat surprising that our *gcn4*Δ mutant was unable to repress the activation of our set of signature genes that we see up regulated when guanine nucleotide generation is inhibited by MPA. Although it (*gcn4*Δ) did further repress the expression of these transcripts in double mutant strains with either of the Bas transcription factors, the *gcn4*Δ single mutant was not able to attenuate hyper-virulence in our murine disseminated infection model as demonstrated by the *bas1*Δ. We have thus shown that Bas1 tightly coordinates the regulation of these pathways in the context of limiting NA conditions even more so than Gcn4. Knowing the involvement of Gcn4 in limiting guanine and in purine biosynthesis it is speculative to consider the Gcn4/TOR (target of rapamycin) pathway in mediating the hyper-virulent

phenotype. One possible scenario that can be proposed follows AICAR accumulation activation of AMPK (AMP-activated protein kinase), which in turn inhibits TORC1. A more direct approach could involve the small GTPase Rheb that positively regulates TORC1 in the GTP-bound state.[77] The potential involvement of TOR is consistent with its role as an energy and nutrient sensor that coordinates anabolic processes and cell cycle progression.

The failure of our Bas1 conditional knockout likely needs to be revisited with some fine-tuning of the initial conditions prior to mouse infection. Although growth curve tests and protein verifications appeared to indicate a positive outcome for the transient ablation of Bas1, it might be necessary to maintain auxin in the SC growth media for a much longer period of time ahead of transferring the cells into mice. Perhaps Bas1 expression necessitates being completely off from the start of culture growth until cells are harvested for injection. This should not result in a slow growth effect since we have not recorded this (*bas1*Δ) effect at least in our tests of liquid cultures.

Our novel finding shows that the inhibition of IMP-Dehydrogenase elicits the same activation of our signature series of up regulated genes representing the three biosynthetic pathways similar to the limiting NA condition. Equally important is the fact that this gene induction can be

blunted or compensated for, by an excess of guanine supplementation. The inability to convert IMP to GMP also results in hyper-colonization that mimics the previously observed NA limiting phenotype. The key to linking NAD^+ and GMP would be to see if the –NA response can be abrogated by exogenous guanine. Several attempts were made with mixed results therefore it remains to be established if these two conditions can be reconciled and complemented.

Our studies into the virulence mechanisms of *C. glabrata* pathogenesis highlight the importance of the purine biosynthetic pathways and may have important physiological implications for human health regarding neurological deficits and neuronal maintenance. Increased accumulation of SAICAR metabolite derivatives have been reported to be involved in neurological abnormalities (mental retardation and autistic features).[78] Similarly, accumulation of AICAR derivatives have been reported to be associated with severe neurological defects.[79] Since low [NAD^+] leads to increased flux through the purine pathway and ensuing SAICAR/AICAR build-up, it is possible that NAD^+ depletion can be yet another early surrogate indicator for neurological impairment. Reduced levels of [NAD^+] in various tissue cell types associated with aging are of real concern, since this provides an opportunistic environment for the establishment and expansion of *C. glabrata* infections. Recent reports have found that increased NAD^+ levels can

stimulate PGC1α (peroxisome proliferator-activated receptor gamma-coactivator 1α) production in the brain.[80] This was found in mouse models of Alzheimer's disease, which is associated with reduced levels of PGC1α suggesting a protective role for NAD^+ while simultaneously inhibiting yeast infections. One striking report links NAD^+ depletion and compromised levels of ATP and GTP with the devastating neurological disease Lesch-Nyhan – which presents with degrees of severity for delayed motor development, renal failure and/or gout along with a very dramatic compulsion of self-mutilation in children.[81,82] The condition arises due to an aberrant HGPRT (hypoxanthine-guanine phosphoribosyltransferase) enzyme, part of the 'salvage pathway' for converting purine bases to nucleotides. The highest activity of HGPRT is found in the brain. In yeast, loss of this gene leads to deregulation of the purine *de novo* pathway.[83] Additionally, guanine based purines are known to be involved in synaptic transmission through the action of GEFs and GAPs (**G**uanine **e**xchange **f**actors and **G**TPase **a**ctivating **p**roteins) modulators of heterotrimeric G proteins associated at the cellular membrane that regulate many second messenger cascades. Other consequences related to guanine depletion may affect energy metabolism and modulation of glutamatergic neurotransmission among other more familiar roles in gene replication and expression.

It remains to be seen whether the prevention of the final reaction products GMP or AMP by disruptions of the pathway downstream of the SAICAR and AICAR signals will all lead to a hyper-colonizing phenotype. Thus far, we've established that inhibition of the conversion of the first committed step (IMPD) in GMP synthesis indeed leads to a hyper-virulent phenotype. We have also confirmed that this phenotype is typically characterized by the activation of three biosynthetic pathways (histidine, purine, one-carbon metabolism) whether by MPA or by low [NAD^+]. Taken together our findings indicate that an imbalance of nucleotide levels of NAD^+ or GMP impacts several cellular functions (e.g. Redox reactions, protein deacetylation - HDACs) and signaling systems (eg. GEFs/GAPs) that in *C. glabrata* convergently result in a pathogenic phenotype. Previous work from the Cormack lab established the exploitation of NAD^+ auxotrophy and the derepression of adhesin proteins. Our current work extends these findings to show a direct consequence of NAD^+ depletion on purine biosynthesis. The loss of GMP generation through the absence of the NAD^+ cofactor or by IMPD inhibition leads to an imbalance of nucleotide pools and a series of potentially unknown downstream effects due to [guanine] limitation. In *C. glabrata* infections this leads to kidney hyper-colonization (UTI) and an overall yeast burden. The result is likely not only due to a proliferation of the infectious organism but is also coupled to a set of survival mechanisms activated by low NAD^+ and compromised nucleotide levels that result in increased adherence

accompanied by an inability of immunological clearance by the human host. Implications for nucleotide imbalances in human diseases accords continued investigation in yeast as a model organism to study the effects of proper maintenance of the nucleotide pools NAD^+, GMP, AMP and their impingement on the central nervous system.

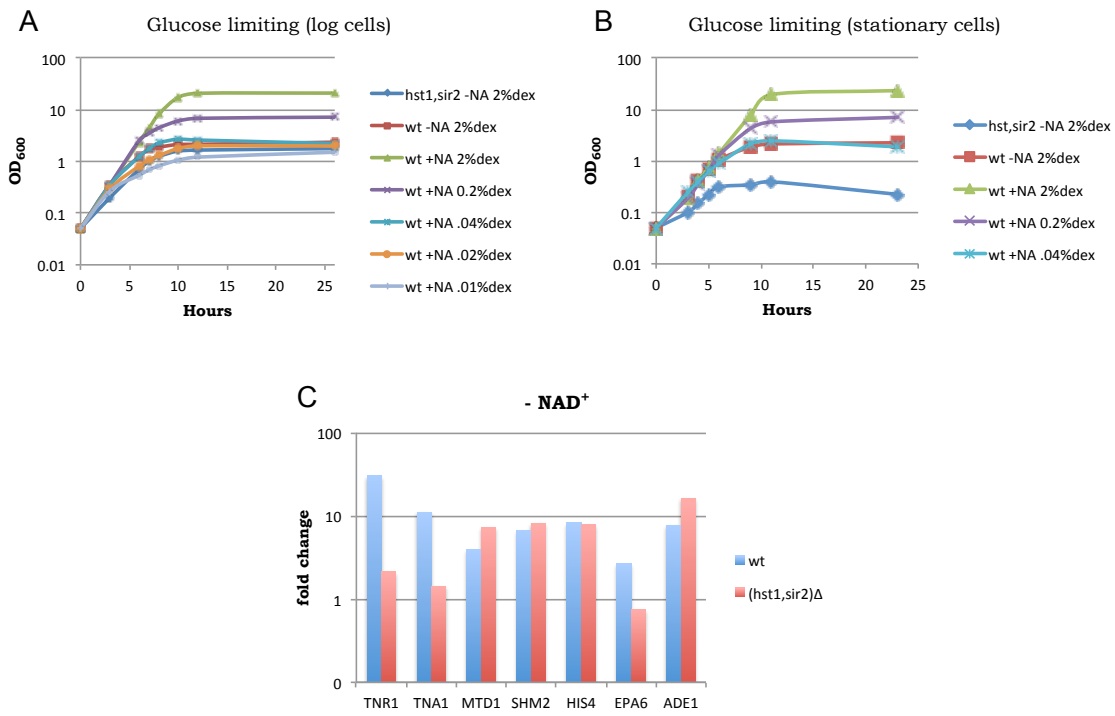


Figure 3.1 Growth curves of BG2-wt in a range of glucose concentrations and (hst1,sir2)Δ mutant analysis.

A, B. *C. glabrata* cells of #2781 wild type and #1764 *hst1*Δ,*sir2*Δ (lab stocks) respectively, were grown overnight in SDC +NA media to both log phase and stationary phase then diluted to an OD_{600} of 0.05 and inoculated into media containing varying dextrose concentrations (0.01, 0.02, 0.04, 0.2, and 2%). The results established an OD matched +NA control by limiting for glucose.

C. qRT-PCR Analysis for (*hst1,sir2*)Δ exposed to –NA conditions. Shown are the nicotinic acid transporters (*TNR1, TNA1*) and representative genes of three biosynthetic pathways (purine, histidine and single carbon THF – tetrahydrofolate).

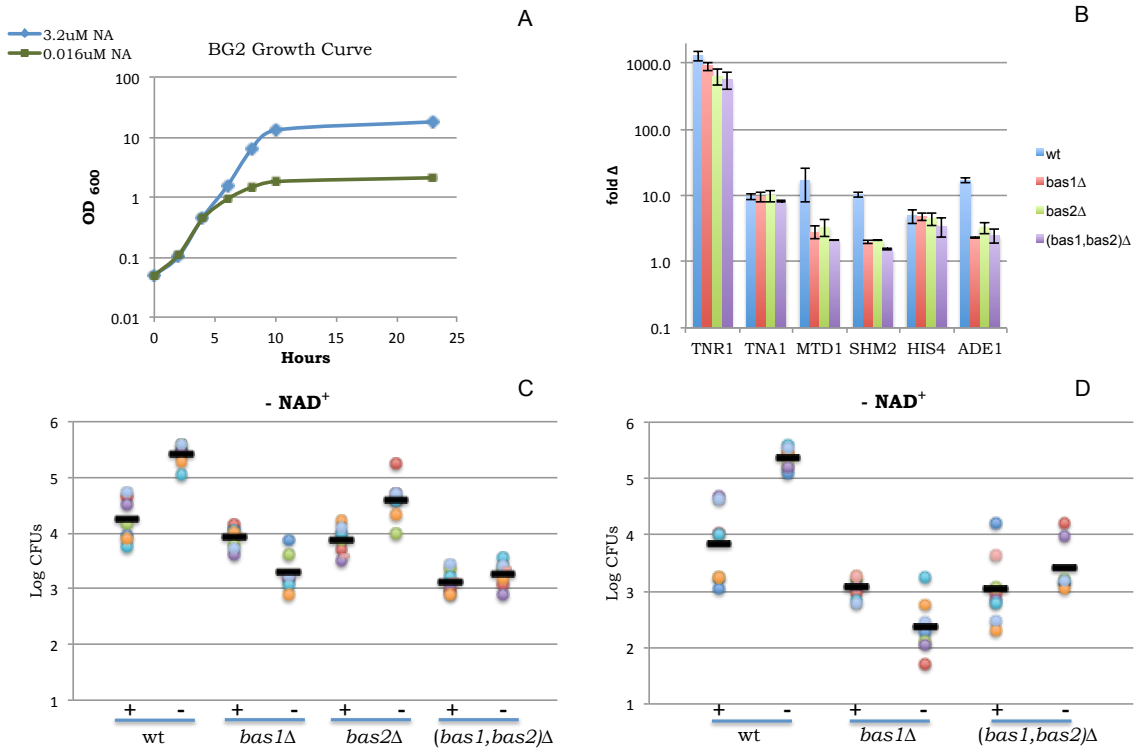


Figure 3.2 qRT-PCR Expression Analysis for BAS Gene Deletion Strains and BAS Mutants in Mouse Disseminated Infections.

A. Low growth effects under limiting NA conditions in SDC media for BG2-wt. B. qRT-PCR results for the inductions of representative genes of purine biosynthesis pathways including the NA transporters as controls. Cells harvested at log phase. Shown are the averages of three experiments that indicate loss of gene expression in the *BAS*Δ mutant strains.

C and D are the results from two separate mouse infection experiments that show *BAS1*Δ strain appears to have an attenuating effect on *C. glabrata* hyper-virulence. *BAS2*Δ may have a marginal effect (Experimental details in Material & Methods sections).

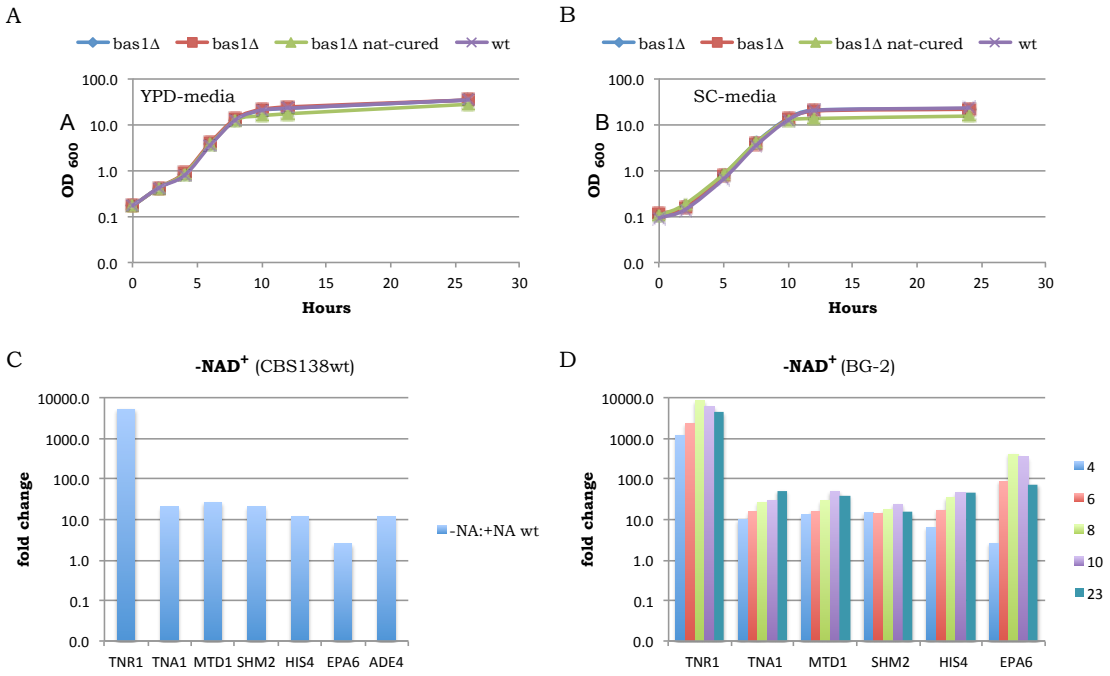


Figure 3.3 BAS1Δ Growth Curves and qRT-PCR analysis of –NA Gene Induction.

A and B. Growth curves for several *bas1*Δ isolates and strains including BG2-wt control. From single colonies, cells (*bas1*Δ: #2521, #2522 isolate#1 and isolate#2 respectively, *bas1*Δ nat-cured #3354, BG2 #2781) were grown overnight to stationary phase in both YPD and SDC+NA (both supplemented with Adenine) then diluted to OD_{600} of 0.1 for 24hr growth monitoring. C. qRT-PCR results of the limiting NA effect of *C. glabrata* strain (CBS138). D. qRT-PCR gene expression monitoring of –NA response.

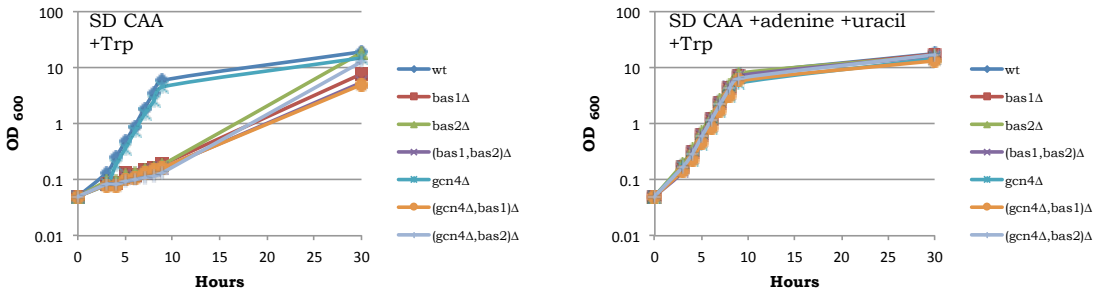


Figure 3.4 Media Tests for Optimal Growth Conditions of Mutant Δ strains.

The *gcn4*Δ strain shows no particular growth phenotype even when grown in casamino acid (CAA) media as long as it is supplemented with Tryptophan. The remaining BASΔ strains or combinations with BAS mutants display expected slow growth phenotypes in the absence of Adenine.

Strains as ordered in the figure correspond to *C. glabrata* lab stock #s:

2781 = BG2- wild type

2521 = Bas1Δ

2523 = Bas2Δ

1947 = (Bas1,Bas2)Δ

3260 = Gcn4Δ

3310 = (Gcn4Δ,Bas1)Δ

3311 = (Gcn4Δ,Bas2)Δ

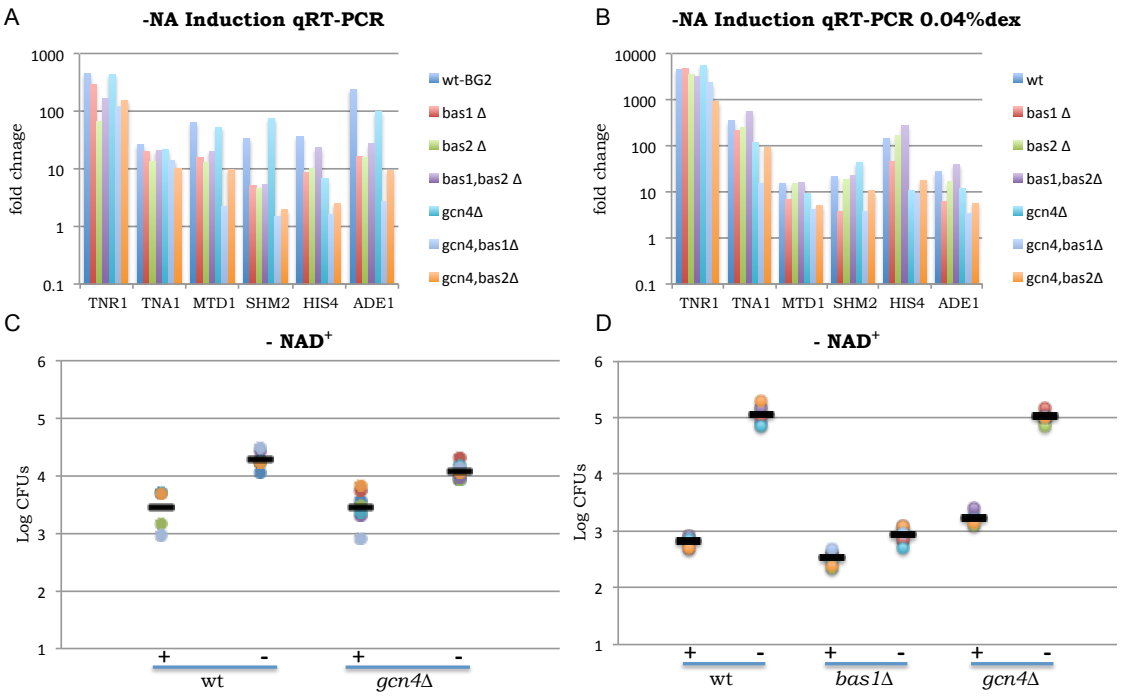


Figure 3.5 GCN4Δ Mutant in Mouse Disseminated Infections and qRT-PCR analysis of BAS and GCN4 mutants under –NA conditions.

qRT-PCRs for the various deletion strains show an additive suppressive effect between the *BAS*Δ's and *GCN4*Δ. A. expression changes comparing deeply –NA starved cells to early +NA log phase cells. Panel B shows similar results while the +NA cells grown in limiting 0.04% glucose were similarly OD matched to –NA starved cells.

Panels C and D are the results from two separate mouse infection experiments showing *BAS1*Δ having an attenuating effect on hyper-virulence while the *GCN4*Δ plays no role in the hyper-colonization. This is a second isolate for the *bas1*Δ mutant showing reproducible and consistent effects. The strains BG2-wt, *bas1*Δ and *gcn4*Δ correspond to

strains #2781, 2522 and 3260 respectively. In panel D the +NA cells grown in limiting 0.02% glucose. (Experimental details in Material & Methods sections)

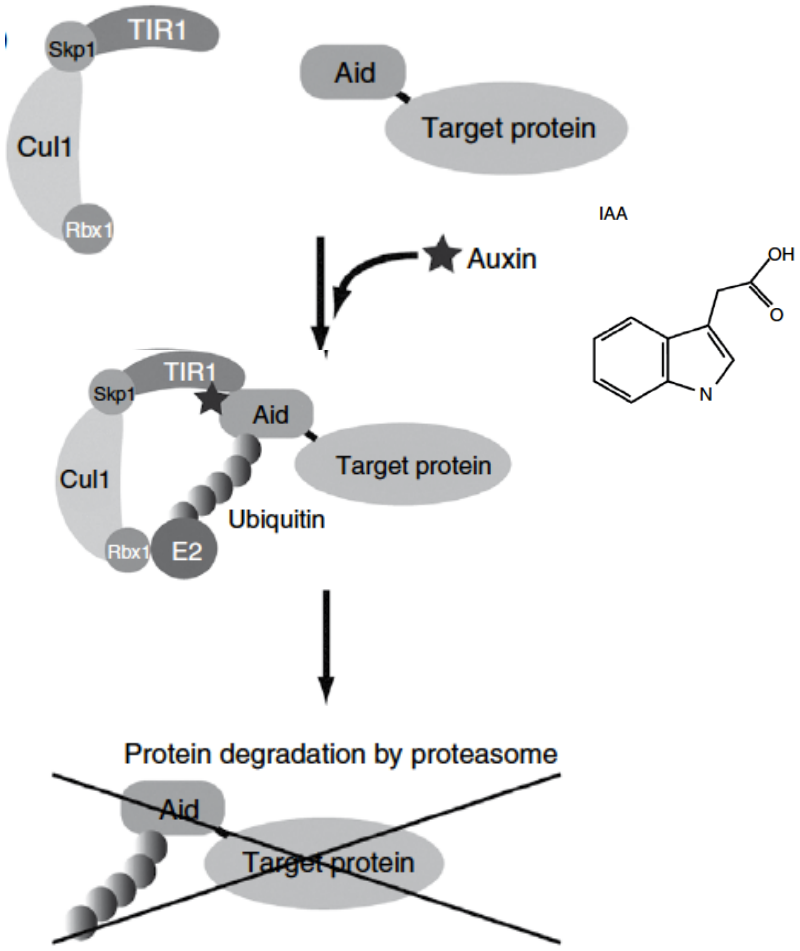


Figure 3.6 Schematic of the AID system and indole-3-acetic acid (Auxin).

Auxin binding to TIR1 promotes the interaction between TIR1 and the AID - degron of the target protein. SCF-TIR1 acts as an E3 ubiquitin ligase to recruit an E2 ligase resulting in polyubiquitylation of the aid degron and targeted destruction by proteasome.

(An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature Methods Vol.6 No12, December 2009, 917)

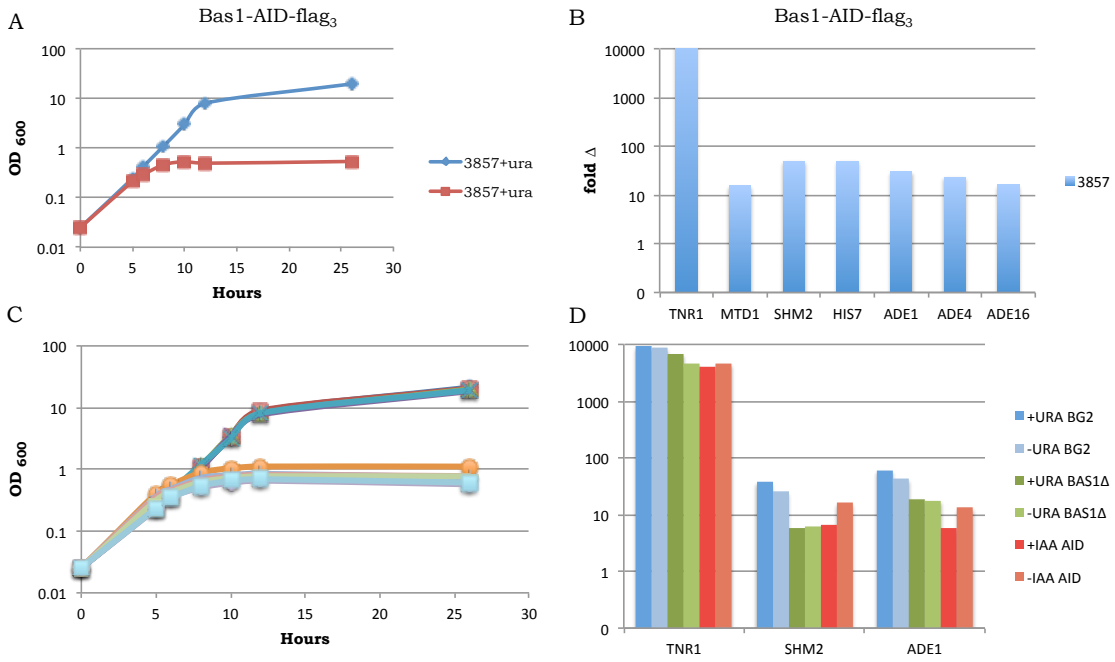


Figure 3.7 Bas1 Conditional Mutant tests with qRT-PCR under –NA conditions.

A. Shows the growth curve for Bas1 – AID – $flag_3$ (#3857) strain grown in –NA limiting conditions (SDC+URA). B. qRT-PCR results showing the induction of representative pathway genes indicating Bas1 – AID – $flag_3$ is functional.

C. Shows the growth curves for BG2 (#2781), Bas1Δ(#2521), Bas1-AID-$flag_3$ TIR1 (#3859). The TIR1 (3859) component is maintained in URA-SDC media while Bas1 is degraded in the presence of Auxin (IAA); added at the start of culture growth under –NA conditions. Panel D shows the induction of representative genes under limiting NA and indicates the

loss of Bas1 by IAA and suppression of gene expression that is similar or worse than *bas1*Δ levels.

Both sets of cells for expression analysis were harvested at 12hrs –NA vs. log phase 6hrs +NA.

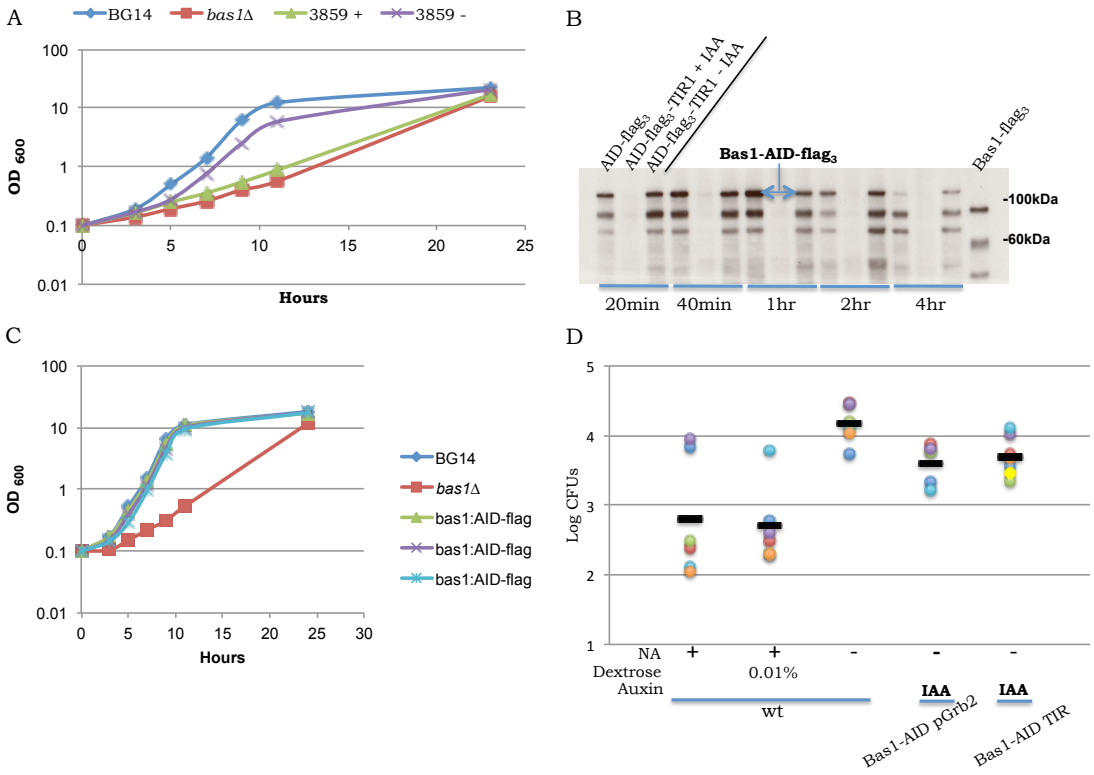


Figure 3.8 Bas1 Conditional Mutant – AID Auxin test in-vitro and Mouse Disseminated Infections (UTI).

A. Growth Curve of the various Bas1 constructs in SC +NA –ADE and SC +NA –ADE –URA for the TIR1myc strain:

#2521 – Bas1 Δ

#3859 – Bas1 AID flag, TIR1myc **+IAA**

#3859 – Bas1 AID flag, TIR1myc **-IAA**

Growth curves for BG2 (#403), Bas1Δ(#2521), Bas1-AID-flag$_3$ TIR1 (#3859). The TIR1 (3859) component is maintained in URA- SDC media

while Bas1 is degraded in the presence of Auxin (IAA); added at the start of culture grown under –NA conditions.

B. Western analysis of Bas1 degradation time course. Cells were harvested at multiple time points. The system is extremely sensitive as there is no detectable protein by the earliest 20min time point. Bas1 detection is by anti-flag Ab.

C. Growth curve functional test of Bas1-AID-$flag_3$ in SDC+NA –ADE media with 3 isolates #3857.

D. Mouse disseminated infection of *C. glabrata* CFUs (kidney). Limiting for glucose control matches the OD of NA starved cells. Only the –NA sample leads to hyper-colonization. The Bas1 conditional knockout shows a marginal effect and is further confounded by the control empty plasmid.

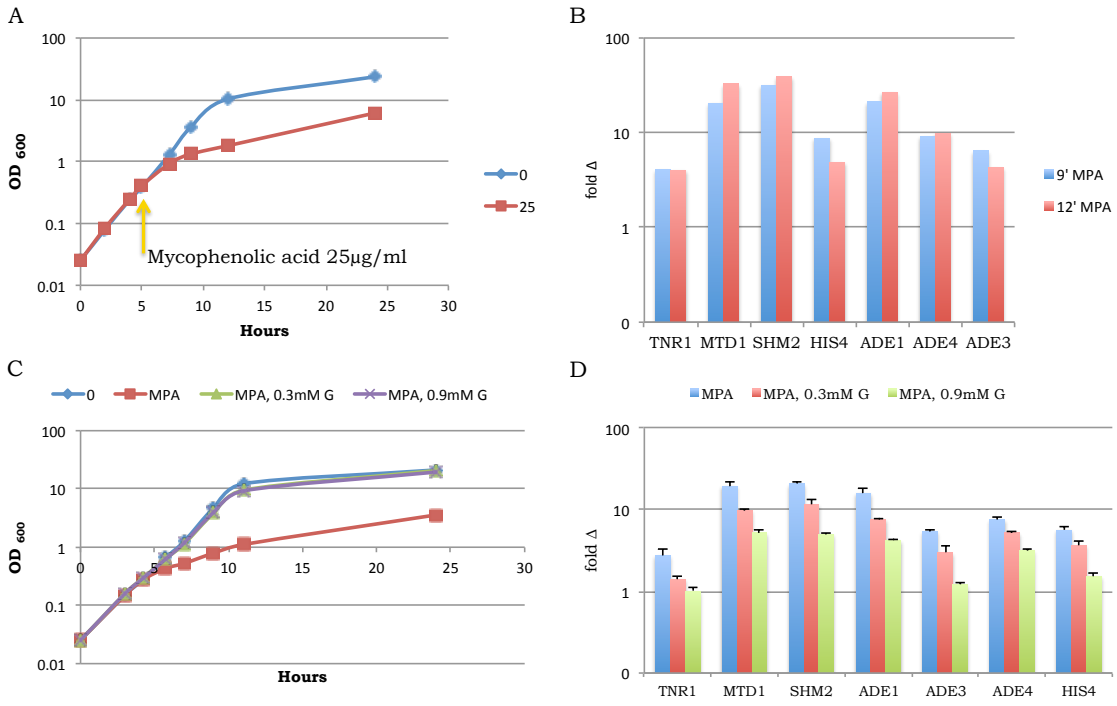


Figure 3.9 IMPD Inhibition. qRT-PCR Induction by MPA and Guanine recovery.

A. Growth curve for BG2-wt #2781 and the affect of the IMPD inhibitor Mycophenolic acid at 25 µg/ml (MPA). B. Shows qRT-PCR result of the induction ratios of genes after +MPA addition vs. no MPA; cell harvesting at hour 9 and 12 growth monitoring. Induction levels are comparable to – NAD^+ levels. C. Growth curve shows that MPA (50 µg/ml) inhibition can be overcome by addition of excess guanine (0.3mM and 0.9mM). D. Shows the qRT-PCR results of induction ratios where gene expression by MPA can be suppressed by addition of guanine almost in a linear fashion down to about basal levels.

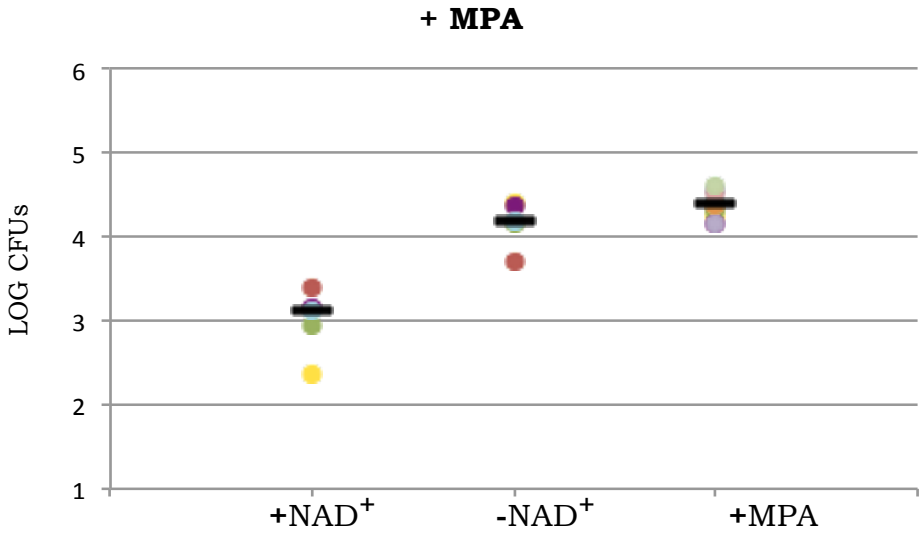


Figure 3.10 IMP-Dehydrogenase Inhibition in Mouse Disseminated Infection (UTI).

C. glabrata cells grown in NAD^+ replete conditions (+), NAD^+ limiting conditions (-), or NAD^+ replete + MPA (50 µg/ml) prior to infection. GMP/GTP starvation replicates NAD^+ limiting hyper-colonization phenotype.

	WT -NA : +NA	(*hst1,sir2*)Δ -NA : +NA
TNR1	332	3
TNR2	245	2
TNA1	10	0
EPA6	21	2
ADE1	6	5
ADE6	6	3
ADE4	8	7
ADE5/7	2	1
ADE3	2	2
ADE17	2	2
ADE8	3	3
HIS4	20	36
HIS2	3	3
HIS7	7	13
HIS1	7	11
MTD1	3	2
SHM2	6	6
GCV1	2	2
GCV3	2	2
SHM1	2	1

Table 3.1 (hst1, sir2)Δ Microarray Summary.

The genes in this table represent the signature set of responders to limiting NA conditions. Included are the NA transporter genes recognized to be under Hst1, Sir2 regulation (blue highlighted values). The remaining sets of genes are confirmed to be up-regulated similar to wild type. (Values are $\log_2$ transformed)

	+NA wt:*bas1Δ*	-NA wt:*bas1Δ*	WT -NA : +NA	*bas1Δ* -NA : +NA
ADE1	1.9	11.0	11.9	2.4
ADE6	1.9	6.6	4.4	1.9
ADE4	2.2	4.3	5.5	2.6
ADE5/7	2.7	5.1	3.9	1.7
ADE13	1.2	7.0	3.8	1.5
ADE3	1.6	2.5	2.8	2.6
ADE17	1.3	2.8	3.2	1.7
ADE8	1.2	3.9	3.2	1.2
HIS4	1.9	2.4	6.3	7.4
HIS2	1.1	1.0	5.6	3.1
HIS7	1.3	1.3	4.0	4.1
HIS1	1.1	1.2	3.4	5.5
MTD1	5.5	8.6	7.0	7.7
SHM2	1.6	7.5	9.2	2.4
GCV1	2.1	2.7	5.0	5.9
GCV3	1.5	2.7	4.0	2.7
SHM1	1.2	1.7	1.7	1.5

Table 3.2 BAS1Δ Microarray Analysis.

The genes assembled in the table are part of a set of conspicuously up-regulated members of three biosynthetic pathways (adenine regulon, histidine and single carbon THF – tetrahydrofolate). A summary of log transformed values representing the average fold changes for two

microarray experiments. Blue highlighted numbers are notable differences in gene expression not expected in the comparisons.

Name	Sequence	Enzyme
bas1 5end 5'	agtggatccGAATCTTCGGGATCTACTAGATC	BamHI
3'bas1_flag	tatcgtcgtcatctttgtaatccttgtcatcgtcgtccttgtagtc ACTTGGGTTAAGAGGAAGAAA	Fusion PCR
5'flag(12)auxin	cgatgatgataaaAAAGAAAAGAGTGCTTGCCC AAAA	Fusion PCR
3'auxin_XmaI	cccgggTCATTTATACATTCTCAAGTC	XmaI
5' CTA1_XmaI	GAATGTATAAATGAcccgggGAAGTTCCTATA CTTTCTAGAG	XmaI
3'CTA1_NheI	gttgagatgagaaaagctagcgct	NheI
bas1 3end 5'	gtcaagcttAATCTTATTCATTTACTATACTGCT C	HindIII
bas1 3end 3'	actgtcgacTGAAAATGAGTTCACGAACTACTC	SalI
5'upAIDflag	AAAAATAGACTTGAGAATGTATAAAGGAGCAGG TGCTgactacaaggacgacgatga**caaggattacaaagatga cg**	Fusion PCR
3'dnAIDflag	CTCTAGAAAGTATAGGAACTTCTCAtttatcatcatcg tccttatagtccttatcgt**cgtcatctttgtaatccttg**	Fusion PCR
5'TIR1_exo	CAATTGCCAAAAAACATTAACATATGACGTACTT CCCGGAGGAGG	Gibson
3'TIR1_exo	TGGGTTGTGTTctcgaggTTAGCTAGTGGATCCGT TCAAGTCT	Gibson

Table 3.3 BAS1-AID-Flag$_3$ Oligo Sequences.

The BAS1 AID knockout construct was made with multiple fusion PCRs and through a combination of site directed cloning and Gibson exo-III cloning.

(Intended to be blank)

Chapter 4 Materials and Methods

4.1 Secondary Structure Prediction and EPA Protein Sequence Alignments

The secondary structure prediction of *C. glabrata* EPA1 ORF was conducted by the open resource PredictProtein (PP; https://www.predictprotein.org). The service provides a breadth of protein structural annotations including secondary structure features such as disulfide bridges, inter-residue contacts, transmembrane helices, beta-barrel structures and solvent accessibility. The purpose for the use of this tool was to enhance the probability of creating a series of structurally stable Epa1 NT fragments encompassing the NT (ATG-K350) region as defined by the lectin domain (Rigden, et.al). The usage of individual amino acids at the +2 position is one important factor in determining protein stability and expression level.

The EPA family of proteins (25) was aligned using the open resource T-Coffee (http://tcoffee.crg.cat). T-Coffee is a multiple sequence alignment server used to create the multiple sequence alignment of Epa proteins to the PA-14 domain – more precisely, the experimentally identified Epa 1 NT lectin binding region (S40 – I283). This homologous region defined the

construct boundaries for small-scale mammalian expressions and eventual glycan array analysis. Details of constructs found on Table 4.

4.2 Expression of Epa1 in *E.coli*

Initial heterologous expression and characterization of the recombinant ligand binding domains of multiple Epa proteins were focused on developing an *E. coli* system that will allow for efficient production of Epa proteins prior to alternative expression systems. I engineered and carried out initial characterizations for several recombinant Epa1 N-terminal domain constructs using an inducible bacterial expression system for purification. Affinity tagged (6xHis) Epa1 was expressed as a fusion to sumo (small ubiquitin-like modifier) using a modified pET-28a vector (Novagen) retaining most of its features including the Kanamycin marker. Pertinent features are detailed in Figure 3. Fourteen fragments corresponding to the lectin domain and differing in their precise N- and C-termini as described in Table 2 were created as *BamHI – XhoI* PCR products. These were cloned into the modified (His-sumo) pET-28a plasmid for expression and purification of fusion constructs. Sequenced verified fragments were transformed into BL-21 Rosetta 2(DE3) competent cells, for "universal" translation of eukaryotic codons in *E. coli*. Rosetta2 cells use a Chloramphenicol resistant plasmid to express rare tRNA genes. Small starter cultures (5ml LB +Cam+Kan) were grown at 37°C and used to inoculate scaled-up cultures at ~ 1:100 dilutions.

0.5mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) was used to induce Epa expression at an OD_{600} of 0.6-0.8. Cells were chilled briefly on ice (5min) and transferred from 37°C to an 18°C incubator shaking at ~225 rpm and grown overnight. The His-tag along with sumo (Smt3) can be removed by Ulp1 (ubiquitin-like protein) specific protease. This step is typically carried out directly on the IMAC (immobilized metal-ion affinity chromatography) column or post elution followed by a second pass over Ni^{2+}. Initial purification was carried out using an IMAC, Ni^{2+}-NTA (Qiagen) column or by batch purification. Cells were spun down at 5000 x *g*, 5min and pellets were resuspended in 0.5M NaCl, 10mM Tris – pH 8.0, 3mM $CaCl_2$ and lysed with 10 µg/mL lysozyme. Lysates were incubated with Ni^{2+} agarose resin in batch for at least 1 hour or passed over a 10mL resin column by gravity flow for large-scale preparations. Bound material was washed with 10 CV (column volumes) 150mM NaCl, 10mM Tris-8.0, 3mM $CaCl_2$ and eluted in 4-5 CV 250mM Imidazole, 150mM NaCl, 10mM Tris-8.0, 3mM $CaCl_2$. Samples ultimately run on SDS-PAGE for Coomassie stain identification. (Protease Inhibitor Cocktail tablets used for lysate buffers were supplied and prepared according to (ROCHE) manufacturer instructions. LB Luria-Bertani antibiotic media and agar plates used the following working concentrations: Kan^R 20 µg/mL, Cam^R 34 µg/mL).

4.3 Cell Culture Maintenance

Suspension adapted HEK-293 GnTI$^-$ cells were originally obtained courtesy of Dr. Reeves (MIT, Boston, MA.). The cells lack N-acetyl-glucosaminyltransferase I (GnTI$^-$) activity, and consequently are unable to synthesize complex N-glycans. HEK-293 GnTI$^-$ cells were maintained in Freestyle media (Invitrogen) and supplemented with 1% FBS, 2mM glutamine and 100 µg/ml (penicillin/streptomycin) and passaged at no less than 2.0 x 10^5 cell density.

4.4 Adapting GnTI$^-$ HEK Cells (Dan Leahy)

Cells were removed from a 95% confluent T25plate by gently banging the flask in the presence of 1-3 day old media (5 ml 10% FBS DMEM:F12/ Pen/strep). Cells were transferred in old media to a small shaker bottle containing 10 ml Freestyle 293 / 1% FBS/added Glutamine 2mM final/Pen/strep). Cells should be around 0.2 x 10^6 per ml at 3 days count, split by spinning cells (1000rpm, 5 minute, 4°C), and resuspend cells in fresh Freestyle media (+1% FBS, + glutamine, keep cells under 2 x10^6/ml) to about 0.1 to 0.2 X 10^6/ml. The most important part of adapting these cells is to vortex in the fresh media each time you split until you see 100% single cells when you are ready to passage again. Expect that to be at least 3-5 passages. Once the cells are all single,

scale up to larger volumes/cell number. Cells will grow only to 1-2 million/mL.

4.5 Mass Spectrometry

Epa 1 protein samples were prepared for mass spec analysis by the Johns Hopkins Mass Spectrometry and Proteomics Facility as detailed by the submission protocols. Briefly, protein samples were run on NuPAGE Bis-Tris 4-12% gels and stained with Coomassie Blue. Bands were cut precisely and pieces were placed in sterile eppendorf tubes. Samples were then washed with 50% EtOH and submitted for analysis immediately.

4.6 Expression and Purification of Epa Gene Family

The Epa1, 6 and 7 N-terminal domains were refined and exactly determined by limited proteolysis experiments and mass spec analysis. The region (S40-I283) was used as a guiding template for identifying homologous domains among all the 25 EPA gene family members. Protein sequence alignments were made using the open resource T-Coffee as described above. Corresponding Epa N-term domain gene sequences were PCR amplified from *C. glabrata* (BG2), and cloned as *BamHI*/*XhoI* or compatible *BglII*/*SalI* fragments into the mammalian expression plasmid pαSHS (Cormack-lab, JHU), a modified version of

pαSHP-H (adapted from Dan Leahy, JHU, Baltimore, MD.) Constructs were expressed as soluble, secreted domains with an amino terminal 6xHis tag along with a sumo fusion tag. Details of the individual sequence constructs can be found in reference Table 4. Purification protocols were similarly carried out as detailed earlier in the bacterial expression system, with minor modifications. The proteins were made as secreted fusion fragments obviating the need for cell lysing. After Ni^{2+} purification, a second affinity-binding step to a lactose-sepharose column was performed. Elution fractions from Ni^{2+} binding were combined and passed over a 0.5mL lac-column by gravity flow. The column was washed with 20 CV of 150mM NaCl, 10mM Tris-8.0, 3mM $CaCl_2$ and eluted with 4-5 CV of 200mM β-lactose. Samples were concentrated on Amicon-ultra (10K MW) spin down columns at 3500 rpm for ~ 10min. Epa proteins were confirmed by Coomassie stain and/or Western analysis. The use of LacNac-sepharose (a substrate for Epa1 binding) column was a final step used to help determine which family members are stable and potential functional lectin binding proteins.

4.7 Mammalian HEK293 GNTI$^-$ Transfection

Cells were transfected at 1-2 $x10^6$ cell density with 100 μg of plasmid DNA and 300 μg of PEI (polyethylenimine, Polysciences Inc.), incubated for 15 minutes in HSFM (Hybridoma Serum Free Media, Invitrogen) and added directly to cells. Several different DNA:PEI ratios were explored,

ultimately a 1:2 or 1:3 ratio worked optimally. Recombinant Epa proteins were harvested 3 days post transfection from 100ml suspension cultures and purified in batch using (Qiagen) Ni^{2+}-NTA resin. Bound material was washed 1x with 20 CV (column volumes) of Binding Buffer (BB- 0.5M NaCl, 20mM Tris-pH8.0, 1mM $CaCl_2$). Consecutive 20CV washes were performed with BB + 20mM Imidazole and BB + 50mM Imidazole. Purified protein was eluted with BB + 250mM Imidazole and verified by Western blotting with an anti-His antibody (Santa Cruz, SC-2837). Subsequently, proteins were promptly dialyzed into 150mM NaCl, 20mM Tris-8.0 and 2mM $CaCl_2$ for use directly onto glycan microarrays.

4.8 Preparation of PEI (Polyethylenimine, Polysciences Inc. cat# 23966)

When making DNA complexes, add in DNA in a drop wise manner to PEI mixing carefully by pipetting. Mix and leave for 15-20 min before adding to cells. Complexes in SFM (serum free media) should be 10% of total final volume.

1. Add 50 mg of PEI to a beaker, adding approximately 40 ml of deionized water.
2. While stirring add 6 N HCl dropwise to the solution until it is ~ pH 2.0.
3. Stir the PEI solution until it is dissolved, maintaining the pH at ~ 2.0 with HCl.

4. Add 6 N NaOH dropwise until the solution reaches pH 7.0.
5. Adjust the final volume to 50ml with deionized water.
6. Pass the solution through a 0.22-micron filter.
7. Aliquot to desired volumes using aseptic technique.
8. Store the aliquots at –20°C or –70°C.

4.9 Western Analysis

Protein samples were separated according to their molecular weights using sodium dodecyl sulfate denaturing polyacrylamide gel electrophoresis (SDS-PAGE) followed by western analysis using standard protocols. Briefly, a broad detection range of 4-12% gradient NuPAGE Bis-Tris gels (Invitrogen) were used for all protein work. Epa proteins were first transferred onto a (Polyvinylidene Fluoride) PVDF membrane using Invitrogen's XCell II™ Blot Module, blocked in 1xPBS - 5% low fat dry milk for at least one hour or incubated overnight at 4°C. Subsequently, the primary antibody in 3% milk 1xPBS was incubated for at least one hour. The antibodies used were specific for either the poly-histidine tag (Santa Cruz #8065) or the N terminal of Epa 1 (lab stock #3638 rabbit polyclonal), also cross-reactive against Epa6 and Epa7 as well. After 1xPBS washes (3x), an HRP linked secondary antibody in 1xPBS was incubated for at least 30 min. Antibody working dilutions were determined empirically but usually 1:5K or 1:10K worked well. After a final 1xPBS, wash (3x), the blot was incubated with the substrate from

(enzyme linked chemi-luminescence) ECL Plus Western Blotting Detection Reagents (GE Healthcare) according to manufacturer protocols. The blot was quickly covered in saran-wrap and placed in an autoradiography cassette for film exposure.

4.10 RNA Extraction for qRT-PCR and Microarrays

The following protocol originally from the Corden lab, was modified by Brian Green in our lab. Briefly, yeast cells (20-30 OD units) were harvested on 0.45µm filters through a filtration unit and washed with 4°C H_2O 2x. Cells were then recovered in H_2O and placed into 1.5 ml screwcap tube. Cells were centrifuged for 30s in a microfuge, the supernatant was removed, and the pellet flash frozen in liquid nitrogen and stored at -80 C before preparing the RNA. Pellets were placed in 350 µl AE buffer (0.85 ml 3 M NaOAc pH5.2, 1 ml 0.5M EDTA, 48.15 ml DEPC H_2O), 50 µl 10% SDS, and 400 µl acid phenol, and incubated at 65°C for 1h in Eppendorf shaker at 900 rpm. Every 15 m, the samples were vortexed, then on ice for 5m, spun for 5 m in microfuge max speed at 4°C. The aqueous layer was transferred to new tube with 400 µl acid phenol and vortexed, then incubated at 65°C for 5m with shaking and vortexed again and back on ice for 5 m. Then, spun for 5m in microfuge at max speed at 4°C. The aqueous layer was placed in new tube with 400 µl chloroform and vortexed and placed on ice for 5m. Spin again for 5m in microfuge at max speed at 4°C. Aqueous layer was placed in new tube

with 40 µl 3M sodium acetate pH5.2 and 1ml 100% ethanol, then placed in -80 C freezer for 20m to 1h. Tubes were spun for 15m in microfuge at max speed at RT. The supernatant was removed and 1 ml RT 70% ethanol made with DEPC H_20 was added. Spin again for 5m in microfuge at max speed at RT then pour off supernatant, quick spin, and aspirate liquid completely with pipette. Let tube sit open for 10 m on ice to dry pellet. We usually continue with a DNAse step in DEPC H_2O and another phenol: chloroform extraction and 70% ethanol wash. RNA is then diluted 1:50 and run on an agarose gel for quality control.

4.11 Glycan Microarrays Analysis

Glycan array analysis was performed by the Consortium for Functional Glycomics; protocol CFG#2556 (http://www.functionalglycomics.org). Lectin proteins were isolated and purified as described earlier by a two-step protocol (Ni2+ and Lac-sepharose) followed by fluorescent labeling. The proteins were conjugated with the Alexa-Fluor 488 dye for detection. Resulting data was sorted and tallied in Excel into major catalogue groups; linkage type, terminal sugar, hydroxyl C-positions, penultimate sugars and glycan modifications to generate relative binding figures for each lectin tested.

4.12 DNA Microarrays In-house

For each sample, total RNA was prepared (as described above) by warm acid phenol extractions and cDNA was made by first strand synthesis (SuperScript II, Invitrogen#11904018) or homemade RT. Visual inspection of the 28S and 18S rRNAs on agarose gels was used to ensure quality of RNA isolation. Total RNA was DNAse treated and resuspended in DEPC-H_2O before cDNA production. Finally, RNA was spectroscopically quantified at 260 nm. The integrity of RNA was verified by an optical density (OD) absorption ratio OD_{260}/OD_{280} between 1.8 and 2.0. The following is a brief overview of the array setup; a Li lab protocol modified by Brian Green in our lab. Cy3 or Cy5 conjugated cDNA was prepared for each sample using anchor dT primers and a 3mM aa-dUTP/dNTP mix (4:1 aa-dUTP:dTTP). Differentially labeled experimental and control cDNA preparations were pooled, heated to 100°C for 2 min, and hybridized overnight at 63°C in 3xSSC, 25mM HEPES pH7.0, 0.25% SDS on in-house oligonucleotide microarrays. Arrays were washed 3x by vigorous plunging in 2xSSC, 0.2% SDS at 60°C, rinsed 1x in 0.1x SSC, spun dry and scanned with a GenePix 4000B (Axon Instruments). Spots were found and analyzed with GenePix Pro 3.0 analysis software. Image data files of the GenePix Results (GPR) were imported into the Acuity database for analysis. Ratio-based normalization of arrays is the method implemented by GenePix Pro. Log ratios have the advantage of being

normally distributed and can easily reveal properties of the data. The software also applied filters for background and spot similarities so that low intensity spots at or near background levels and spots that differed more than twofold in their expression ratio were eliminated. We note specifically that we did not obtain sufficient reproducible replicates to properly analyze our data through SAM (significance of analysis of microarrays) software. Our tables were generated by hand for a number of chosen genes whose values were averaged after $\log_2$ transformations.

4.13 Quantitative RT-PCR

A number of strains were used for these studies:

Strain #	Gene
2781	BG2-wild type
2521	*bas1*Δ
2523	*bas2*Δ
1947	*(bas1,bas2)*Δ
1764	*(hst1,hst2)*Δ
3260	*gcn4*Δ
3310	*(gcn4,bas1)*Δ
3311	*(gcn4,bas2)*Δ

The strains were grown in SDC (synthetic complete amino acids) and excess NA at 30°C overnight to log phase (OD_{600} ~0.5), washed in PBS 2x and placed in fresh media at OD_{600} ~ 0.05 with replete (3.25 μM NA) or in the absence of NA. Cells were recovered at log phase or during stationary phase according to experimental design. RNA was extracted as described

above except our first strand synthesis is done with homemade RT. 1 µl of cDNA was used as the template in real time PCR reactions with primer pairs for *TNR1, TNA1, MTD1, SHM2, HIS4, ADE1, ADE3, ADE4, EPA6* and *HHT2*. Real time PCR was performed on Bio-Rad C1000 Thermal Cycler in 96-well plates. Real-time PCR efficiencies were acquired by serial dilutions of 10ng μl^{-1} of genomic *C. glabrata* DNA to generate standard curves for each gene. RNA equivalents were normalized to *HHT2* transcript levels. qRT-PCRs were done in duplicate wells from each of at least three independent biological samples for experiments that display error bars. A master-mix (20 µl $well^{-1}$) of the following reaction components was prepared: 8.25 µl of water, 2.5 µl of 10x PCR buffer, 2.5 µl of Primer mix (or 1.25 µl of Forward, 1.25 µl of Reverse), 2.5 µl of 10mM $MgCl_2$, 2.5 µl of dNTP mix (2 mM each), 0.25 µl of *Taq* polymerase, 1.25 µl of EvaGreen 20x stock (Molecular Probes, Eugene, Oregon), aliquoted into each of the wells on the 96-well PCR plate. 5 µl cDNA per well was used; from 1:20 dilution. The following general real-time PCR protocol was used: Heat activation of polymerase (in house-Taq) for 2 min at 95°C, 40 cycles of PCR: 95°C for 10 s, 65°C for 20 s, 72°C for 20s. Melting curve analysis was carried out to confirm the specificity of the amplicons. Quantification of gene expression differences was performed using the 'delta-delta-CT' method with HHT2 as the normalizer gene.

4.14 Animal Studies

Nine-week old Balb/c female mice (Taconic) were typically used in our limiting NA (nicotinic acid) experiments. Prior to tail vein injections *C. glabrata* strains were grown at 30°C in SDC (synthetic complete + 2% dextrose). When cells reached mid-log phase (OD: 0.55) they were diluted back and shifted to SDC media in the presence of NA 3.25 μM or absence of NA and grown to stationary phase. Cells were collected by centrifugation and washed 2x in ice cold PBS and resuspended at a concentration of 1 x 10^7 cells ml^{-1}. Numbers of cells were determined by spectrophotometry OD_{600}. Mice were injected with ~1 x 10^7 cells in PBS via tail vein injections in a total volume of 0.1ml. Each yeast strain used a group of 7 to 10 mice per experiment. Mice were sacrificed and target organs were harvested seven days post-infection. Tissue homogenates were plated on YPD plates supplemented with Pen/Strep (Penicillin 100U ml^{-1}, Streptomycin 100 μg ml^{-1}). Viability and counts were confirmed by assessing cfu from appropriate dilutions on YPD Pen/Strep plates, incubated for 24 -48 hrs at 30°C. CFUs for all samples were tallied and the geometric means for all groups were determined.

4.15 Cloning of Bas1 – Auxin Inducible Degron (AID)

The AID system was generated in a series of steps with multiple fusion PCRs into the pOZ16 URA^+ plasmid. The yeast strain #3857 was created

by integration of the BAS1- $FLAG_3$-AID-$FLAG_3$ construct with a URA^+ marker and a hygromycin selection cassette into the parental strain BG14 $URA^=$ (#430), done by a two-step method. The hyg^R marker would be removed by transformation with pRD16 FLP (flippase). Schematically, the auxin inducible degron (AID) tag was subcloned from the Auxin plasmid BYP7430 and placed immediately 3' of the BAS1 ORF. This was followed by 3x-Flag tag and a stop codon. *BAS1* 500bp fragments corresponding to the terminal ORF region and the 3' UTR genomic region immediately past the stop sequence were PCR amplified and cloned into pOZ16 as *BamHI – NheI* and *HindIII – SalI* fragments respectively. The BamHI fragment was made with two fusion PCRs consisting of Fusion I (flag-AID 700bp and 3'CTA1 200bp) product. Fusion II was made with BAS1-flag 500bp and the Fusion I product. The construct is now BAS1-flag-AID in pOZ16. A Flag tag (3x D-Y-K-D-D-D-D-K) was incorporated into the appropriate oligos following *BAS1* and AID. The second flag tag was made as a fusion PCR consisting of the BAS1-flag-AID-**flag** (1.2kb) and **flag**-CTA1 (250bp) products. The overlapping flag sequence is denoted in bold. The BAS1 *HindIII – SalI* fragment was placed 3' of the FRT site downstream of the Hyg^R marker. The TIR1 component was introduced into pGrb2.3 (B2039) at *XbaI – XhoI* sites as a Gibson cloning PCR product (New England Biolabs #E5510S).

A noteworthy caveat: when generating AID tagged proteins, never place the AID degron exposed at the C-terminal end as this will result in immediate degradation of your desired protein.

4.16 Gene Deletions in *C. glabrata*

Gene deletion in *C. glabrata* was done by a one-step method. The ORF of the gene was either replaced with a nourseothricin (NAT) cassette or completely deleted. The sequence information was obtained from the Genolevures website as well as our own in-house sequencing database of the BG2 strain. The genes *BAS1*, *BAS2,* (*BAS1,2*) and *GCN4* were deleted in this manner. For deletion with NAT cassette, the 5' and 3' untranslated regions of each gene to be deleted, as well as NAT cassette were PCR amplified. The three PCR fragments were then used as templates for fusion PCR where the three pieces were joined into one. The fusion PCR product was purified with the Qiagen gel purification kit then transformed into *C. glabrata* via the lithium acetate method and selected on YPD plates supplemented with 200 μg ml^{-1} NAT. Homologous recombination and allele replacement of each locus was verified by PCR analysis using a primer that anneals in the sequences external to the cloned fragments and a primer annealing within the NAT cassette. We also verified the absence of a gene by inability to PCR amplify an internal fragment from each deleted gene.

4.17 Strains and Media

Yeast cells were grown on standard yeast media YPD or synthetic complete (SC-Hopkins) supplemented with 2% dextrose. When necessary SDC media was supplemented with Adenine 25 µg/ml or Tryptophan 85 µg/ml. All *Candida glabrata* deletion strains were derived from our wild type lab strain BG2 or from an *ura3* derivative of BG2, BG14. All bacteria were grown in LB medium supplemented with 100 µg ml^{-1} carbenicillin at 37°C. For plates LB + agar was added. When needed plates were also supplemented with 500 µg ml^{-1} Hygromycin or 100 µg ml^{-1} Carbenicillin.

Chapter 5 Conclusion

These two major projects were aimed at understanding the virulent pathways of the yeast pathogen *C. glabrata*. In the first study we addressed a structural biology question by creating a unique system to generate a functional protein of a critical Epa domain known to be responsible for the binding of yeast cells to human host epithelial cells. Understanding the essence of the binding interface in the first step of pathogenesis is essential toward developing new therapies against fungal infections. Additionally, a series of glycan microarrays experiments led to further support and characterize the ligand binding specificities of a family of lectin adhesion proteins known as epithelial adhesins (EPA) to host glycan moieties. The second study focused on delineating the regulatory pathways that are activated during exposure to an NA limiting environment that lead to a hyper-colonization phenotype.

For the first project, we fell short of solving the X-ray crystallographic structure of the Epa1 although those efforts were refocused toward the entire family of adhesins that yielded a wealth of glycan binding information for six different EPA adhesins. It was an effort that entailed the creation of 25 different expression constructs and generated a series of Epa ligand binding domains for functional testing. For Epa1, 6 and 7 our data were consistent with the importance of key cysteine residues for

the functional architecture of the PA-14 domain within the N-terminal region. This idea was borne out in the published structure of the Epa1p. Additionally, our work showed that for a number of these lectins (Epas 12, 23, and 25), they exclusively prefer binding to sulfated glycan residues. Sulfated residues are found in a number of different tissues and are a preferred strategy for gaining access and colony establishment for many other disease causing microorganisms. The high degree of binding specificity of the EPA proteins for varying types of cell surface glycans present a tempting opportunity for utilizing these proteins as non-invasive diagnostic strategies for detecting aberrant glycosylation underlying various disease states.

The second major study addressed the hyper-colonization phenotype observed during a mouse model of disseminated infection of *C. glabrata*. This virulence program is characterized by the up regulation of a set of genes that control three distinct pathways: purine and histidine biosynthesis and 1-Carbon metabolism. By making use of a series of deletion mutants in genes orthologous to the *S. cerevisiae BAS1* and *BAS2* genes, genes known to be involved in the regulation of these central metabolic pathways in *S. cerevisiae*, we could show that Bas1 is required for induction of these metabolic pathways in response to a low nicotinic acid (NA) signal. A low niacin environment is the critical inducer of the pathogenic state. Analysis of these mutants through microarrays

and qPCR experiments alerted to the presence of additional factors necessary for complete induction or complementary to the *bas1* knockout, and *GCN4*, a regulator of amino acid biosynthesis genes in *S. cerevisiae* was also uncovered as an indirect participant in the activation of several genes in purine biosynthesis independent of Bas1p. It was discovered that although *GCN4* is not required for the hyper-colonization phenotype, it did in fact have a synergistic affect with *BAS1* resulting in a complete failure of gene expression when yeast cells were exposed to limiting NA conditions. Direct cooperative association between Bas1p and Gcn4p remains to be established.

Why were these three pathways so strongly up regulated? Additional sets of experiments attempted to answer that question and tested the downstream affects of the absence of NA as an enzymatic cofactor that directly impacts the generation of GMP/GTP. We hypothesize that the absence of NA leads to an inhibition of IMP to GMP conversion. In support of this hypothesis, we were able to mimic the limiting NA phenotype in our mouse infection experiments via inhibition of IMP-dehydrogenase with mycophenolic acid, an inhibitor of IMPDH. These results helped take another step forward in identifying how limiting NA alters virulence of *C. glabrata*.

Intended to be blank

References

1. Glöckner A, Cornely OA. Candida glabrata - unique features and challenges in the clinical management of invasive infections. *Mycoses*. 2015;58(8):445-450. doi:10.1111/myc.12348.

2. Fidel PL, Vazquez JA, Sobel JD. Candida glabrata: Review of Epidemiology, Pathogenesis, and Clinical Disease with Comparison to C. albicans. *Clin Microbiol Rev*. 1999;12(1):80-96. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC88907/.

3. Lepak AJ, Marchillo K, Andes DR. Pharmacodynamic Target Evaluation of a Novel Oral Glucan Synthase Inhibitor , SCY-078 (MK-3118), Using an In Vivo Murine Invasive. 2015;59(2):1265-1272. doi:10.1128/AAC.04445-14.

4. Weinstein RA, Lundstrom T, Sobel J. Nosocomial Candiduria: A Review. *Clin Infect Dis* . 2001;32(11):1602-1607. doi:10.1086/320531.

5. Cormack BP, Ghori N, Falkow S. An adhesin of the yeast pathogen Candida glabrata mediating adherence to human epithelial cells. *Science*. 1999;285(5427):578-582. doi:10.1126/science.285.5427.578.

6. Kobayashi O, Hayashi N, Kuroki R, Sone H. Region of Flo1 proteins responsible for sugar recognition. *J Bacteriol*. 1998;180(24):6503-6510.

7. Frieman MB, Cormack BP. The -site sequence of glycosylphosphatidylinositol-anchored proteins in Saccharomyces cerevisiae can determine distribution between the membrane and the cell wall. *Mol Microbiol*. 2003;50(3):883-896. doi:10.1046/j.1365-2958.2003.03722.x.

8. Frieman MB, Cormack BP. Multiple sequence signals determine the distribution of glycosylphosphatidylinositol proteins between the plasma membrane and cell wall in Saccharomyces cerevisiae. *Microbiology*. 2004;150(Pt 10):3105-3114. doi:10.1099/mic.0.27420-0.

9. Weig M, Jänsch L, Gross U, De Koster CG, Klis FM, De Groot PWJ. Systematic identification in silico of covalently bound cell wall proteins and analysis of protein-polysaccharide linkages of the human pathogen Candida glabrata. *Microbiology*. 2004;150(Pt 10):3129-3144. doi:10.1099/mic.0.27256-0.

10. Frieman MB, McCaffery JM, Cormack BP. Modular domain structure in the Candida glabrata adhesin Epa1p, a β1,6 glucan-cross-linked cell wall protein. *Mol Microbiol*. 2002;46(2):479-492. doi:10.1046/j.1365-2958.2002.03166.x.

11. Rigden DJ, Mello L V., Galperin MY. The PA14 domain, a conserved all-?? domain in bacterial toxins, enzymes, adhesins and signaling molecules. *Trends Biochem Sci*. 2004;29(7):335-339. doi:10.1016/j.tibs.2004.05.002.

12. Petosa C, Collier RJ, Klimpel KR, Leppla SH, Liddington RC. Crystal structure of the anthrax toxin protective antigen. *Nature*. 1997;385(6619):833-838. http://dx.doi.org/10.1038/385833a0.

13. Castaño I, Pan S-J, Zupancic M, Hennequin C, Dujon B, Cormack BP. Telomere length control and transcriptional regulation of subtelomeric adhesins in Candida glabrata. *Mol Microbiol*. 2005;55(4):1246-1258. doi:10.1111/j.1365-2958.2004.04465.x.

14. Iraqui I, Garcia-Sanchez S, Aubert S, et al. The Yak1p kinase controls expression of adhesins and biofilm formation in Candida glabrata in a Sir4p-dependent pathway. *Mol Microbiol*. 2005;55(4):1259-1271. doi:10.1111/j.1365-2958.2004.04475.x.

15. De Las Peñas A, Pan S-JJ, Castaño I, et al. Virulence-related surface glycoproteins in the yeast pathogen Candida glabrata are encoded in subtelomeric clusters and subject to RAP1- and SIR-dependent transcriptional silencing. *Genes Dev*. 2003;17(18):2245-2258. doi:10.1101/gad.1121003.

16. Domergue R. Nicotinic Acid Limitation Regulates Silencing of Candida Adhesins During UTI. *Science (80-)*. 2005;308(5723):866-870. doi:10.1126/science.1108640.

17. Kaur R, Domergue R, Zupancic ML, Cormack BP. A yeast by any other name: Candida glabrata and its interaction with the host. *Curr Opin Microbiol*. 2005;8(4):378-384. doi:http://dx.doi.org/10.1016/j.mib.2005.06.012.

18. de Groot PWJ, Bader O, de Boer AD, Weig M, Chauhan N. Adhesins in human fungal pathogens: Glue with plenty of stick. *Eukaryot Cell*. 2013;12(4):470-481. doi:10.1128/EC.00364-12.

19. Zupancic ML, Frieman M, Smith D, Alvarez RA, Cummings RD, Cormack BP. Glycan microarray analysis of Candida glabrata adhesin ligand specificity. *Mol Microbiol*. 2008;68(3):547-559. doi:10.1111/j.1365-2958.2008.06184.x.

20. De Groot PWJ, Kraneveld EA, Qing YY, et al. The cell wall of the human pathogen Candida glabrata: Differential incorporation of novel adhesin-like wall proteins. *Eukaryot Cell*. 2008;7(11):1951-1964. doi:10.1128/EC.00284-08.

21. Gräslund S, Sagemark J, Berglund H, et al. The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. *Protein Expr Purif*. 2008;58(2):210-221. doi:10.1016/j.pep.2007.11.008.

22. Bivona L, Zou Z, Stutzman N, Sun PD. Influence of the Second Amino Acid on Recombinant Protein Expression. 2011;74(2):248-256. doi:10.1016/j.pep.2010.06.005.Influence.

23. Varshavsky a. The N-end rule: functions, mysteries, uses. *Proc Natl Acad Sci U S A*. 1996;93(October):12142-12149. doi:10.1073/pnas.93.22.12142.

24. Ielasi FS, Decanniere K, Willaert RG. The epithelial adhesin 1 (Epa1p) from the human-pathogenic yeast Candida glabrata: Structural and functional study of the carbohydrate-binding domain. *Acta Crystallogr Sect D Biol Crystallogr*. 2012;68(3):210-217. doi:10.1107/S0907444911054898.

25. Blixt O, Head S, Mondala T, et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci United States Am* . 2004;101(49):17033-17038. doi:10.1073/pnas.0407902101.

26. Maestre-reyna M, Diderrich R, Stefan M, Eulenburg G, Kalugin V, Brückner S. Structural basis for promiscuity and speci fi city during Candida glabrata invasion of host epithelia. 2012:1-6. doi:10.1073/pnas.1207653109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1207653109.

27. Veelders M, Brückner S, Ott D, Unverzagt C, Mösch H-U, Essen L-O. Structural basis of flocculin-mediated social behavior in yeast. *Proc Natl Acad Sci U S A*. 2010;107(52):22511-22516. doi:10.1073/pnas.1013210108.

28. Galili U. Significance of the Evolutionary ??1,3-Galactosyltransferase (GGTA1) Gene Inactivation in Preventing Extinction of Apes and Old World Monkeys. *J Mol Evol*. 2014;80(1):1-9. doi:10.1007/s00239-014-9652-x.

29. Galili U. Anti-Gal: An abundant human natural antibody of multiple pathogeneses and clinical benefits. *Immunology*. 2013;140(1):1-11. doi:10.1111/imm.12110.

30. Rosen SD. Ligands for L-Selectin: Homing, Inflammation, and Beyond. *Annu Rev Immunol*. 2004;22(1):129-156. doi:10.1146/annurev.immunol.21.090501.080131.

31. Rosen SD. Endothelial ligands for L-selectin: from lymphocyte recirculation to allograft rejection. *Am J Pathol*. 1999;155(4):1013-1020. doi:10.1016/S0002-9440(10)65201-7.

32. Martini R, Xin Y, Schmitz B, Schachner M. The L2/HNK-1 Carbohydrate Epitope is Involved in the Preferential Outgrowth of Motor Neurons on Ventral Roots and Motor Nerves. *Eur J Neurosci*. 1992;4(7):628-639. doi:10.1111/j.1460-9568.1992.tb00171.x.

33. Pomin VH. Sulfated glycans in inflammation. *Eur J Med Chem*. 2015;92:353-369. doi:10.1016/j.ejmech.2015.01.002.

34. Fontsere N, Cardozo C, Donate J, et al. Tunneled Catheters with Taurolidine-Citrate-Heparin Lock Solution Significantly Improve the Inflammatory Profile of Hemodialysis Patients. *Antimicrob Agents Chemother*. 2014;58(7):4180-4184. doi:10.1128/AAC.02421-14.

35. Achkar JM, Fries BC. Candida infections of the genitourinary tract. *Clin Microbiol Rev*. 2010;23(2):253-273. doi:10.1128/CMR.00076-09.

36. Kojic EM, Darouiche RO. Candida Infections of Medical Devices. *Clin Microbiol Rev*. 2004;17(2):255-267. doi:10.1128/CMR.17.2.255-267.2004.

37. Recuenco FC, Kobayashi K, Ishiwa A, et al. Gellan sulfate inhibits Plasmodium falciparum growth and invasion of red blood cells in vitro. *Sci Rep*. 2014;4:4723.

38. Ishiwa A, Kobayashi K, Takemae H, et al. Effects of dextran sulfates on the acute infection and growth stages of Toxoplasma gondii. *Parasitol Res*. 2013;112(12):4169-4176.

39. Yanagishita M, Hascall VC. Cell surface heparan sulfate proteoglycans. *J Biol Chem* . 1992;267(14):9451-9454. http://www.jbc.org/content/267/14/9451.short.

40. Hayman JR, Southern TR, Nash TE. Role of sulfated glycans in adherence of the microsporidian *Encephalitozoon intestinalis* to host cells in vitro. *Infect Immun*. 2005;73(2):841-848. doi:10.1128/IAI.73.2.841-848.2005.

41. Frevert U. Malaria circumsporozoite protein binds to heparan sulfate proteoglycans associated with the surface membrane of hepatocytes. *J Exp Med*. 1993;177(5):1287-1298. doi:10.1084/jem.177.5.1287.

42. Wadström T, Ljungh A. Glycosaminoglycan-binding microbial proteins in tissue adhesion and invasion: key events in microbial pathogenicity. *J Med Microbiol*. 1999;48(3):223-233. doi:10.1099/00222615-48-3-223.

43. Carruthers VB, Håkansson S, Olivia K, Sibley LD, Giddings OK. Toxoplasma gondii Uses Sulfated Proteoglycans for Substrate and Host Cell Attachment Toxoplasma gondii Uses Sulfated Proteoglycans for Substrate and Host Cell Attachment. *Infect Immun*. 2000;68(7):4005-4011. doi:10.1128/IAI.68.7.4005-4011.2000.Updated.

44. Alvarez-Dominguez C, Vazquez-Boland J a, Carrasco-Marin E, Lopez-Mato P, Leyva-Cobian F. Host cell heparan sulfate proteoglycans mediate\nattachment and entry of Listeria monocytogenes,\nand the listerial surface protein ActA is involved\nin heparan sulfate receptor recognition. *Infect Immun*. 1997;65(1):78-88.

45. Pethe K, Aumercier M, Fort E, Gatot C, Locht C, Menozzi FD. Characterization of the heparin-binding site of the mycobacterial heparin-binding hemagglutinin adhesin. *J Biol Chem*. 2000;275(19):14273-14280.

46. Chen Y, Maguire T, Hileman RE, et al. Dengue virus infectivity depends on evelope protein binding to target cell heparan sulfate. *Nat Med*. 1997;3(8):866-871. doi:10.1038/nm0897-866.

47. Jackson T, Ellard FM, Ghazaleh R a, et al. Efficient infection of cells in culture by type O foot-and-mouth disease virus requires binding to cell surface heparan sulfate. *J Virol*. 1996;70(8):5282-5287.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=190485&tool=pmcentrez&rendertype=abstract.

48. Herrera EM, Ming M, Ortega-Barria E, Pereira MEA. Mediation of Trypanosoma cruzi invasion by heparan sulfate receptors on host cells and penetrin counter-receptors on the trypanosomes. *Mol Biochem Parasitol*. 1994;65(1):73-83.

49. Butcher BA, Sklar LA, Seamer LC, Glew RH. Heparin enhances the interaction of infective Leishmania donovani promastigotes with mouse peritoneal macrophages. A fluorescence flow cytometric analysis. *J Immunol*. 1992;148(9):2879-2886.

50. Love DC, Esko JD, Mosser DM. A heparin-binding activity on leishmania amastigotes which mediates adhesion to cellular proteoglycans. *J Cell Biol*. 1993;123(3):759-766.

51. Sim L, Groes M, Olesen K, Henriksen A. Structural and biochemical characterization of the N-terminal domain of flocculin Lg-Flo1p from Saccharomyces pastorianus reveals a unique specificity for phosphorylated mannose. *FEBS J*. 2013;280(4):1073-1083. doi:10.1111/febs.12102.

52. Bogan KL, Brenner C. Nicotinic acid, nicotinamide, and nicotinamide riboside: a molecular evaluation of NAD+ precursor vitamins in human nutrition. *Annu Rev Nutr*. 2008;28:115-130. doi:10.1146/annurev.nutr.28.061807.155443.

53. Belenky P, Bogan KL, Brenner C. NAD+ metabolism in health and disease. *Trends Biochem Sci*. 2007;32(1):12-19. doi:10.1016/j.tibs.2006.11.006.

54. Gazzaniga F, Stebbins R, Chang SZ, McPeek MA, Brenner C. Microbial NAD metabolism: lessons from comparative genomics. *Microbiol Mol Biol Rev*. 2009;73(3):529-41, Table of Contents. doi:10.1128/MMBR.00042-08.

55. Dujon B, Sherman D, Fischer G, et al. Genome evolution in yeasts. *Nature*. 2004;430(6995):35-44. doi:10.1038/nature02579.

56. Brunke S, Hube B. Two unlike cousins: Candida albicans and C.glabrata infection strategies. *Cell Microbiol*. 2013;15(5):701-708. doi:10.1111/cmi.12091.

57. Denu JM. Vitamin B3 and sirtuin function. *Trends Biochem Sci*. 2005;30(9):479-483. doi:10.1016/j.tibs.2005.07.004.

58. Denu JM. Linking chromatin function with metabolic networks: Sir2 family of NAD+-dependent deacetylases. *Trends Biochem Sci*. 2003;28(1):41-48. doi:10.1016/S0968-0004(02)00005-1.

59. Ma B, Pan S-J, Domergue R, et al. High-affinity transporters for NAD+ precursors in Candida glabrata are regulated by Hst1 and induced in response to niacin limitation. *Mol Cell Biol*. 2009;29(15):4067-4079. doi:10.1128/MCB.01461-08.

60. Daignan-Fornier B, Fink GR. Coregulation of purine and histidine biosynthesis by the transcriptional activators BAS1 and BAS2. *Proc Natl Acad Sci U S A*. 1992;89(15):6746-6750. doi:10.1073/pnas.89.15.6746.

61. Gauthier S, Coulpier F, Jourdren L, et al. Co-regulation of yeast purine and phosphate pathways in response to adenylic nucleotide variations. *Mol Microbiol*. 2008;68(6):1583-1594. doi:10.1111/j.1365-2958.2008.06261.x.

62. Subramanian M, Qiao WB, Khanam N, et al. Transcriptional regulation of the one-carbon metabolism regulon in Saccharomyces cerevisiae by Bas1p. *Mol Microbiol*. 2005;57(1):53-69. doi:10.1111/j.1365-2958.2005.04663.x.

63. Arndt KT, Styles C, Fink GR. Multiple global regulators control HIS4 transcription in yeast. *Science (80-)*. 1987;237(4817):874-880. http://science.sciencemag.org/content/237/4817/874.abstract.

64. Devlin C, Tice-Baldwin K, Shore D, Arndt KT. RAP1 is required for BAS1/BAS2- and GCN4-dependent transcription of the yeast HIS4 gene. *Mol Cell Biol*. 1991;11(7):3642-3651. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=361116&tool=pmcentrez&rendertype=abstract.

65. Natarajan K, Meyer MR, Belinda M, et al. Transcriptional Profiling Shows that Gcn4p Is a Master Regulator of Gene Expression during Amino Acid Starvation in Yeast. *Mol Cell Biol*. 2001;21(13):4347-4368. doi:10.1128/MCB.21.13.4347.

66. Hinnebusch AG. TRANSLATIONAL REGULATION OF *GCN4* AND THE GENERAL AMINO ACID CONTROL OF YEAST*. *Annu Rev Microbiol*. 2005;59(1):407-450. doi:10.1146/annurev.micro.59.031805.133833.

67. Tice-Baldwin K, Fink GR, Arndt KT. BAS1 has a Myb motif and activates HIS4 transcription only in combination with BAS2. *Science (80-)*. 1989;246(4932):931-935. doi:10.1126/science.2683089.

68. Springer C, Künzler M, Balmelli T, Braus GH. Amino acid and adenine cross-pathway regulation act through the same 5'- TGACTC-3' motif in the yeast HIS7 promoter. *J Biol Chem*. 1996;271(47):29637-29643. doi:10.1074/jbc.271.47.29637.

69. Joo YJ, Kim JA, Baek JH, et al. Cooperative regulation of ADE3 transcription by Gcn4p and Bas1p in Saccharomyces cerevisiae. *Eukaryot Cell*. 2009;8(8):1268-1277. doi:10.1128/EC.00116-09.

70. Pinson B, Vaur S, Sagot I, Coulpier F, Lemoine S, Daignan-Fornier B. Metabolic intermediates selectively stimulate transcription factor interaction and modulate phosphate and purine pathways. *Genes Dev*. 2009;23(12):1399-1407. doi:10.1101/gad.521809.

71. Denis V, Boucherie H, Monribot C, Daignan-Fornier B. Role of the Myb-like

protein Bas1p in Saccharomyces cerevisiae: A proteome analysis. *Mol Microbiol*. 1998;30(3):557-566. doi:10.1046/j.1365-2958.1998.01087.x.

72. Nishimura K, Fukagawa T, Takisawa H, Kakimoto T, Kanemaki M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods*. 2009;6(12):917-923. doi:10.1017/CBO9781107415324.004.

73. Saint-Marc C, Pinson B, Coulpier F, Jourdren L, Lisova O, Daignan-Fornier B. Phenotypic consequences of purine nucleotide imbalance in Saccharomyces cerevisiae. *Genetics*. 2009;183(2):529-538. doi:10.1534/genetics.109.105858.

74. Rébora K, Desmoucelles C, Borne F, Pinson B, Daignan-Fornier B. Yeast AMP pathway genes respond to adenine through regulated synthesis of a metabolic intermediate. *Mol Cell Biol*. 2001;21(23):7901-7912. doi:10.1128/MCB.21.23.7901-7912.2001.

75. Rébora K, Laloo B, Daignan-Fornier B. Revisiting purine-histidine cross-pathway regulation in Saccharomyces cerevisiae: A central role for a small molecule. *Genetics*. 2005;170(1):61-70. doi:10.1534/genetics.104.039396.

76. Roetzer A, Gabaldón T, Schüller C. From Saccharomyces cerevisiae to Candida glabrata in a few easy steps: Important adaptations for an opportunistic pathogen. *FEMS Microbiol Lett*. 2011;314(1):1-9. doi:10.1111/j.1574-6968.2010.02102.x.

77. Arsham AM, Neufeld TP. Thinking globally and acting locally with TOR. *Curr Opin Cell Biol*. 2006;18(6):589-597. doi:10.1016/j.ceb.2006.09.005.

78. Georges J, Berghe V. an Infantile Autistic Syndrome Characterised By the Presence of Succinylpurines in Body Fluids. *Lancet*. 1984;324(8411):1058-1061. doi:10.1016/S0140-6736(84)91505-8.

79. Marie S, Heron B, Bitoun P, Timmerman T, Van Den Berghe G, Vincent M-F. AICA-ribosiduria: a novel, neurologically devastating inborn error of purine biosynthesis caused by mutation of ATIC. *Am J Hum Genet*. 2004;74(6):1276-1281. doi:10.1086/421475.

80. Chi Y, Sauve A a. Nicotinamide riboside, a trace nutrient in foods, is a vitamin B3 with effects on energy metabolism and neuroprotection. *Curr Opin Clin Nutr Metab Care*. 2013;16(6):657-661. doi:10.1097/MCO.0b013e32836510c0.

81. Fairbanks LD, Jacomelli G, Micheli V, Slade T, Simmonds HA. Severe pyridine nucleotide depletion in fibroblasts from Lesch-Nyhan patients. *Biochem J*. 2002;366(Pt 1):265-272. doi:10.1042/BJ20020148 BJ20020148 [pii].

82. Deutsch SI, Long KD, Rosse RB, Mastropaolo J, Eller J. Hypothesized deficiency of guanine-based purines may contribute to abnormalities of neurodevelopment, neuromodulation, and neurotransmission in Lesch-Nyhan syndrome. *Clin Neuropharmacol*. 2005;28(1):28-37. doi:00002826-200501000-00006 [pii].

83. Guetsova ML, Lecoq K, Daignan-Fornier B. The Isolation and Characterization of Saccharomyces cerevisiae Mutants That Constitutively Express Purine Biosynthetic Genes. *Genetics*. 1997;147(2):383-397. http://www.genetics.org/content/147/2/383.abstract.

Carlos Gomez

cgomez2@jhmi.edu • New York, NY 11102

Education

09/06 – 12/16 Johns Hopkins University School of Medicine, Baltimore MD. Ph.D. Cellular and Molecular Medicine

08/04 - 08/06 Columbia University College of Physicians and Surgeons, New York, NY. M.S. Nutrition

09/96 – 05/01 State University of New York (SUNY) at Stony Brook, NY. B.S. Biochemistry

Professional Experience

09/16 – present **Johns Hopkins Institute for Applied Economics, Global Health and Study of Business Enterprise**
Research Assistant for Professor Steve Hanke
- Valuation analysis in the Healthcare sector using proprietary Hanke-Guttridge DCF model along with Monte Carlo simulations. Utilized long-term "asset turns" trends as well as proxy-statement reviews to generate equity investment memos (REGN, MDCO).

01/16 – 05/16 **T. Rowe Price.** Baltimore, MD
Externship – Equity Research Biotech
- Preparation of equity research report with a technical perspective in the biotechnology sector. Perform company valuation analysis, market analysis with a final investment recommendation (NLNK).

Scientific and Laboratory Experience

05/15 – 11/15 **MedImmune.** Gaithersburg, MD
Internship – Drug Delivery and Device Development group
- Cancer targeting by PLGA-PEG polymer based nanoparticles as a drug delivery system. Engineered a series of NP formulations with targeting ligands by way of click chemistry. Successfully showed initial proof of concept in an *in-vitro* model of ovarian cancer.
- Mechanisms of protein drying – investigated a protein formulation at high concentration in needle-based delivery systems in an effort to identify formulation based strategies for minimizing or preventing drying and system clogging. Developed an experimental set-up for evaluating protein drying under environmentally controlled conditions.

09/06 – 05/15 **Johns Hopkins University, Laboratory of Brendan Cormack.** Baltimore, MD
Graduate Fellow
Thesis: A role for *BAS1* in *Candida glabrata* virulence
- Developed a mammalian expression system for the structural analysis of the EPA family of adhesion proteins by crystallography and glycan microarrays.
- Uncovered a potential role for the transcription factor Bas1 in yeast pathogenesis during exposure to limiting levels of vitamin B3 (niacin); utilized deletion mutants and microarrays.

08/03 – 08/06 **Columbia University, Laboratory of Lawrence Shapiro.** New York, NY
Graduate Fellow
Thesis: Proteomics analysis of lipid-droplet monodispersed membrane proteins
- Managed daily progress of a proteomics project in obesity
- Trained incoming technicians and students in essential laboratory protocols

06/99 – 08/03 **Mount Sinai School of Medicine, Laboratory of Zhen-Qiang Pan.** New York, NY
Research Associate
- Characterized novel class of signaling molecules involved in intracellular neuronal gene activation. Tubby proteins were found to function as heterotrimeric-G protein responsive factors.
- Studied immune-deficiency disorders. Characterization of the functional role of RAG proteins using an extensive panel of Alanine mutants highlighting the structural importance of conserved regions.
- Collaborative roll dissecting the functional domain architecture of the ROC1 ubiquitin ligase.

Carlos Gomez

cgomez2@jhmi.edu • New York, NY 11102

Grants and Awards

03/12 **American Society for Microbiology Conference**. San Francisco
Role of *BAS1* in Regulating Purine Metabolism during NAD^+ Limitation
Gomez C, Ma B, Cormack B.

09/09- 06/11 **Ruth L. Kirschstein National Research Service Award (NRSA)**
$42,000 National Institute of Allergy and Infectious Diseases, National Institutes of Health

10/10 **1st Latin American Protein Society Meeting**. Salta, Argentina.
Toward the *EPA1* Structure of the Yeast Pathogen *Candida glabrata*
Gomez C, Zupancic M, Cormack B.

Publications

Gomez C, Ma B, Cormack B. Bas1 as an attenuator of *C.glabrata* virulence. – *manuscript preparation*.

Vitenshtein A, Charpak-Amikam Y, Yamin R, Bauman Y, Isaacson B, Stein N, Berhani O, Dassa L, Gur C, Glasner A, Gomez C, Ben-Ami R, Osherov N, Cormack BP, Mandelboim O. NK Cell Recognition of Candida glabrata through Binding of NKp46 and NCR1 to Fungal Ligands Epa1, Epa6 and Epa7. Cell Host Microbe. 2016 Oct 12;20: 527-534.

Subramanian V, Rothenberg A, Gomez C, Cohen AW, Garcia A, Bhattacharyya S, Shapiro L, Dolios G, Wang R, Lisanti MP, Brasaemle DL. Perilipin A mediates reversible binding of CGI-58 to lipid droplets in 3T3-L1 adipocytes. J Biol Chem. 2004 Oct 1;279(40):42062-71.

Gomez C, Carroll K, Shapiro L. Tubby Proteins: the plot thickens. Nat Rev Mol Cell Biol. 2004 Jan;5(1):55-63. Review.

Santagata S, Boggon TJ, Baird CL, Gomez CA, Zhao J, Shan WS, Myszka DG, Shapiro L. G-protein signaling through tubby proteins. Science. 2001 Jun 15;292(5524):2041-50.

Gomez CA, Ptaszek LM, Villa A, Bozzi F, Sobacchi C, Brooks EG, Notarangelo LD, Spanopoulou E, Pan ZQ, Vezzoni P, Cortes P, Santagata S. Mutations in Conserved Regions of the Predicted RAG2 Kelch Repeats Block Initiation of V(D)J Recombination and Result in Primary Immunodeficiencies. Mol Cell Biol. 2001 Aug;20(15):5653-64.

Santagata S, Gomez CA, Sobacchi C, Bozzi F, Abinun M, Pasic S, Cortes P, Vezzoni P, Villa A. N-terminal RAG1 frameshift mutations in Omenn's syndrome: internal methionine usage leads to partial V(D)J recombination activity and reveals a fundamental role in vivo for the N-terminal domains. Proc Natl Acad Sci U S A. 2000 Dec 19;97(26):14572-7.

Chen A, Wu K, Fuchs SY, Tan P, Gomez C, Pan ZQ. The conserved RING-H2 finger of ROC1 is required for ubiquitin ligation. J Biol Chem. 2000 May 19;275(20):15432-9.

Wu K, Fuchs SY, Chen A, Tan P, Gomez C, Ronai Z, Pan ZQ. The SCF (HOS/beta-TRCP)-ROC1 E3 ubiquitin ligase utilizes two distinct domains within CUL1 for substrate targeting and ubiquitin ligation. Mol Cell Biol. 2000 Feb;20(4):1382-93.

Tan P, Fuchs SY, Chen A, Wu K, Gomez C, Ronai Z, Pan ZQ. Recruitment of a ROC1-CUL1 ubiquitin ligase by Skp1 and HOS to catalyze the ubiquitination of I kappa B alpha. Mol Cell. 1999 Apr;3(4):527-33.

Skills

Bloomberg Terminal: beginner
Technical: Microarrays, protein purification (FPLC), Cell Flow Cytometry, qRT-PCR, Fluorometry | R
Language: Spanish – fluent

Intended to be blank